DISSERTATION

UTILIZING THE ASSESSMENT OF MOTOR AND PROCESS SKILLS ABILITY MEASURES TO PREDICT LEVEL OF COMMUNITY DEPENDENCE

Submitted by

Brenda K. Merritt

School of Education

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Summer 2007

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

UMI Number: 3279530

INFORMATION TO USERS

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleed-through, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

UMI®

UMI Microform 3279530

Copyright 2007 by ProQuest Information and Learning Company.

All rights reserved. This microform edition is protected against unauthorized copying under Title 17, United States Code.

ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

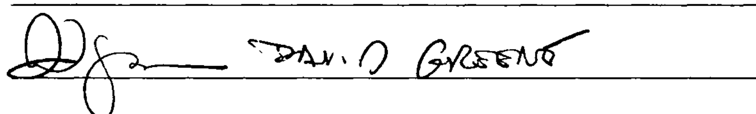
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

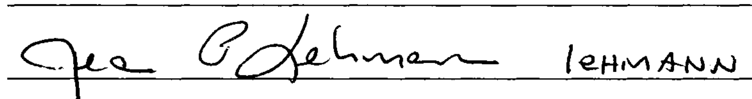
COLORADO STATE UNIVERSITY

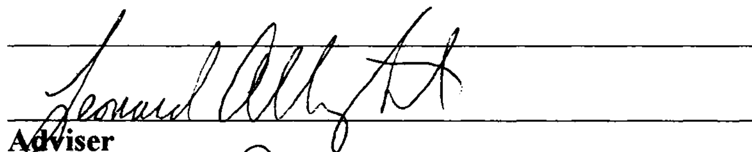
March 9, 2007

WE HEREBY RECOMMEND THAT THE DISSERTATION PREPARED UNDER OUR SUPERVISION BY BRENDA K. MERRITT ENTITLED UTILIZING THE ASSESSMENT OF MOTOR AND PROCESS SKILLS ABILITY MEASURES TO PREDICT LEVEL OF COMMUNITY DEPENDENCE BE ACCEPTED AS FULFILLING IN PART REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY.

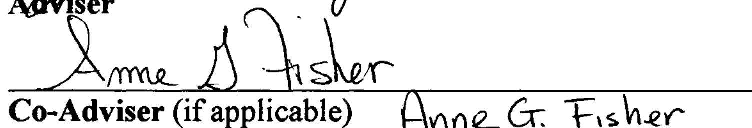
Committee on Graduate Work



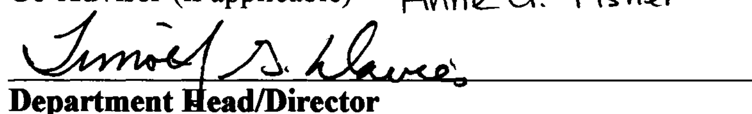




Adviser



Co-Adviser (if applicable) Anne G. Fisher



Department Head/Director

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

ABSTRACT OF DISSERTATION

UTILIZING THE ASSESSMENT OF MOTOR AND PROCESS SKILLS ABILITY MEASURES TO PREDICT LEVEL OF COMMUNITY DEPENDENCE

Objective: To determine if ADL motor and ADL process ability measures can be used to predict level of community dependence.

Participants: Potential participants included all available people in the Assessment of Motor and Process Skills (AMPS) database who (a) had been assessed after January 2000, (b) had not previously been associated with rater scoring error as evidenced by artificially high ADL motor or ADL process ability measures, (c) were not scored by more than 10 raters during rater calibration, and (d) did not have missing or unknown functional level ratings.

Method: Receiver operating characteristic (ROC) curves were generated to (a) verify the accuracy of the current AMPS cutoff measures, (b) determine if AMPS ability measures can be used to determine level of community dependence, and (c) globally examine the relative predictive validity of AMPS ability measures across different diagnostic groups.

Results/Conclusion: Results indicate both the ADL motor and ADL process ability measures can be utilized to help occupational therapists determine level of community dependence. Proposed ADL motor and ADL process cutoff measures demarcating independence were 1.50 and 1.10 logits respectively, and proposed ADL motor and ADL process cutoff measures demarcating the need for moderate to maximal assistance were 1.00 and 0.70 logits respectively. The ability measures between the proposed cutoff measures demarcate a risk zone. For all diagnostic groups the estimates of the area under

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

the ROC curve (global estimates of predictive accuracy) indicated that ADL process ability measure is a fair to good indicator of community dependence. Area under the curve estimates for ADL motor ability were fair to good for all diagnostic groups, except for the psychiatric group, where the area under the curve estimate indicated that ADL motor ability is a poor indicator of community dependence.

Brenda Kathleen Merritt
School of Education
Colorado State University
Fort Collins, CO 80523
Summer 2007

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

TABLE OF CONTENTS

ABSTRACT iii

LIST OF TABLES vii

LIST OF FIGURES ix

CHAPTER 1 – Introduction 1

CHAPTER 2 – LITERATURE REVIEW

Independent Living 17

Modern Measurement and Test Theory 20

Assessments of Activities of Daily Living 34

Assessments of Independent Living 64

Global Judgment of Functional Ability 79

Receiver Operating Characteristic Curves 83

CHAPTER 3 – METHODS

Research Design 96

Participants and Site 97

Instrumentation 99

Procedure/Data Analysis 103

CHAPTER 4 – RESULTS

Decision: Independent versus the Need for Assistance Decision 108

Decision: Other versus Moderate to Maximal Assistance 110

Accuracy of Multiple Cutoff Measures 111

Across Different Diagnostic Categories 112

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

CHAPTER 5 – DISCUSSION

Summary of Findings 114

Interpretation of the Results 118

Clinical Recommendations 124

Objective Measurement 127

Conclusion 129

REFERENCES 131

APPENDICES

Appendix A 176

Appendix B 178

Appendix C 179

Appendix D 180

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

LIST OF TABLES

Table 1 The Assessment of Motor and Process Skills (AMPS) Global Functional Levels 145

Table 2 Tasks Personal Maintenance and Development Domain 146

Table 3 Tasks in the Domain of Homemaking and Community Life 148

Table 4 Tasks in the Domain of Leisure 150

Table 5 Tasks in the Domain of Travel 152

Table 6 Skill Categories and Items in the Assessment of Motor and Process Skills. 154

Table 7 The Functional Independence Measure (FIM) Items and Levels of Scoring 155

Table 8 Items within the Functional Autonomy Measurement System (SMAF) 156

Table 9 Summary Scores for the Katz Index of ADL 157

Table 10 Sample Item from the Klein-Bell ADL Scale 158

Table 11 Global Comparison of ADL Assessment Properties 159

Table 12 Tasks within the Community Living Skills Assessment Inventory 160

Table 13 Scoring Criteria for the Community Living Skills Assessment Inventory 161

Table 14 Items on the St. Louis Inventory of Community Living 162

Table 15 Number of Participants by Diagnostic Group and Global Functional Level 163

Table 16 Mean ADL Motor and ADL Process Ability by Global Functional Level 164

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table 17 ROC Curve Analysis: Using ADL Motor or ADL Process Ability as an Indicator of Community Independence 165

Table 18 Area Under the Receiver Operating Characteristic (ROC) Curve by Decision Condition and Diagnostic Category 166

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

LIST OF FIGURES

Figure 1 Example of a receiver operating characteristic (ROC) curve 167

Figure 2 Identification and definition of the four possible decisions within the current project 168

Figure 3 Impact of raising the cutoff measure or decision threshold 169

Figure 4 Consideration of the relative costs of possible errors 170

Figure 5 ROC curve using ADL motor or ADL process ability to categorize individuals as either independent in the community or in need of assistance to live in the community 171

Figure 6 ROC curve using ADL motor or ADL process ability to categorize individuals either as "other" or in need of moderate to maximal assistance to live in the community 172

Figure 7 Accuracy of correctly categorizing individuals who are independent in the community 173

Figure 8 Accuracy of correctly categorizing individuals who are in need of minimal assistance to live in the community 174

Figure 9 Accuracy of correctly categorizing individuals who are in need of moderate to maximal assistance to live in the community 175

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

CHAPTER 1

INTRODUCTION

The impetus for this research project began in the context of clinical practice after several occupational therapists (including myself) were questioned by other healthcare professionals about the accuracy of our predictions of the amount of assistance that our clients would likely need upon returning to the community. When "pushed" to validate our predictions, we realized that our experience and our clinical reasoning skills were the foundation of our judgments. Providing more formal evidence to document the validity and reliability of our predictions proved to be more difficult. This reflective process led us in search of evidence and a protocol for accurately determining the level of assistance needed for community dwelling. As a member of the staff involved in this process, I felt driven to search for better evidence and methods for predicting the amount of assistance needed to live in the community, also referred to as global functional level and/or level of dependence.

The first step in this investigation involved examining the literature to determine the role occupational therapists have in the evaluation of a client's global functional level. One study reported that occupational therapists were among the four most significant members of the discharge planning team (Gair & Hartery, 2001). In this respect, it is important to consider that the important aspects of the occupational therapy process include (a) evaluating occupational performance (i.e., the client's ability to carry out selected and necessary activities of daily living [ADL] tasks) (Appendix A), (b) planning

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

and implementation of intervention to enhance occupational performance, and (c) evaluating outcomes in terms of ability to perform ADL tasks needed for community living (AOTA, 2002). Thus, during the occupational therapy process, the occupational therapist must identify factors that support or hinder occupational performance, and identify and recommend discharge needs.

An important role of an occupational therapist is to help the rehabilitation and/or healthcare team determine the most appropriate discharge plan. When determining the discharge plan, the healthcare team aims to ensure that the client will be able to safely and efficiently carry out necessary ADL tasks upon discharge. In this process, the occupational therapist helps determine the client's need for services such as (a) caregiver assistance, (b) assistive ADL devices, (c) adaptations to the physical environment, (d) continued therapy, (e) client and/or caregiver education, and/or (f) referral to other agencies and/or healthcare providers.

In determining discharge needs, the occupational therapist considers how much assistance the client will likely need in order to carry out those tasks that are necessary or desired for independent living within the community. Thus, the occupational therapist must consider the client's ability to independently and safely accomplish the assorted daily life tasks within the following areas of independent living: (a) basic self-care (e.g., dressing, grooming, hygiene), (b) food preparation, (c) home maintenance (e.g., cleaning the house, caring for clothes, maintaining the yard), (d) transportation and community mobility, (e) money management, (f) leisure, (g) health and safety (e.g., first aid, emergency procedures, knowledge of proper nutrition, drug/alcohol awareness, safe sex), (h) shopping, (i) social interaction, and (j) self-direction (Agran, Marchand-Martella, &

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Martella, 1994; DeLoach, Wilkins, & Walker, 1983; Dever & Knapczyk, 1988; Globe Fearon, 1997; Webber, Jenkinson, & McGillivray, 2002).

Since determining level of dependence is a role of occupational therapists, the next step in my research involved gaining a better understanding of the means by which occupational therapists accomplish this task. I initiated this step by informally interviewing several experienced occupational therapists, including faculty members at Colorado State University and occupational therapists who were currently working in acute and transitional care facilities. All of the therapists that I interviewed commented that they utilized information gathered through the use of interviews and assessments of occupational performance. Many also said they used assessments of body functions (standardized and nonstandardized), assessments of social and physical environments, and at times, informal assessment of social interaction skills. For clarity, it is important to understand that occupational performance refers to the engagement in and performance of meaningful and necessary daily life tasks (Fisher, 2006c) whereas body functions pertain to the physiological function of the person's body systems (Appendix A) (Fisher, 2006c).

Upon inquiring as to the common methods used to determine level of dependence, there emerged a wide variety of beliefs and preferences among the therapists. Some therapists firmly believed that the assessment of occupational performance was the foundation and cornerstone in helping them determine their client's discharge needs and level of assistance needed to live in the community. Activities of daily living assessments used within the clinic include the evaluation of the client's ability to perform personal ADL (PADL) tasks (e.g., bathing, toileting, dressing, grooming) as well as

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

instrumental ADL (IADL) tasks (e.g., cooking, home maintenance, shopping) (Appendix A). In fact, the occupational therapists in the acute and transitional care settings told me that decisions about level of dependence are made on a daily basis with nearly every client through utilizing assessments of ADL (standardized and nonstandardized) and clinical judgment.

In contrast, a few therapists believed that tests of body function, especially assessments of cognitive function, were also useful and necessary in helping them determine their client's level of dependence. Despite their differences, when I inquired about the evidence behind their assessment choices, each therapist commented that further evidence was needed to verify the validity of their assessments. Thus, from a clinical perspective, there was a recognized need to go beyond expert experience and investigate more deeply the available evidence to support and/or refute the various methods that are currently being utilized by occupational therapists to predict level of dependence.

Although many of the experts that I interviewed said that they supplement their decisions of global functional level by using tests of body functions, researchers have documented that such assessments are not accurate indicators of disability, occupational performance, and/or global functional level (DeBettignies & Mahurin, 1989, Dubuc, Haley, Kooyoomjian, & Jette, 2004; Dickerson, 1997; Gross & Battié, 2006; Mercier, Audet, Hébert, Rochette, & Dubois, 2001; Rice, Leonard, & Carter, 1998; Lindén, Boschian, Eker, Schalén, & Nordström, 2005; Searight & Golderberg, 1991; Winograd, 1984). In fact, in reference to assessing clients with dementia, one author stated, "to date, these studies reveal an inadequacy of mental status measures as predictors of capacity for

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

independent living and underscore the need for including an assessment of independent living skills" (DeBettignies & Mahurin, p. 464). Likewise, within the area of mental health, Dickerson stated "standardized instruments to assess patients' psychotic symptoms . . . do not assess patients' community functioning" (1997, p. 897). Serright and Goldberg (1991) commented that assessments of psychiatric symptoms may not be related to competence within daily life. Lastly, Rice et al. (1998) found that greater physical strength did not indicate better ADL task performance.

In a more recent study, Meinow, Kåreholt, and Lagergren (2005) found that the strongest predictors of amount of assistance received in the home were dependency in PADL ($R^2 = 0.34, p < 0.001$), dependency in IADL ($R^2 = 0.39, p < 0.001$), and cognitive function ($R^2 = 0.13, p < 0.001$). Although cognitive function was among the top three predictors of the needed amount of assistance received, it is important to note that it was a weaker predictor than those related to ADL performance, and only explained 13% of the variance with regard to the amount of assistance needed. The lack of consistent and strong evidence that tests of body function(s) can be used to accurately assess independent living skills and/or predict level of dependence within the community led me to conclude that future occupational therapy research should instead focus on evidence of the use of ADL assessments in determining level of dependence.

A formal review of the literature confirmed that ADL assessments are utilized by occupational therapists to assist in predicting level of dependence. As there is a detailed description of some of ADL assessments often used by occupational therapists in Chapter 2, I present in this section only a brief review of each assessment's concurrent and/or predictive qualities.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

The Assessment of Motor and Process Skills (AMPS) (Fisher, 2006a, 2006b) ADL motor and ADL process ability measures have been shown to correlate significantly with global functional level ($r = .48$, $p < .01$ and $r = .59$, $p < .01$ respectively) (A. G. Fisher, personal communication, January 13, 2005). Additionally, 93% of those with AMPS ADL process ability measures below the cutoff measure of 1.0 logit likely need assistance to live in the community (Fisher, 2006a; Hartman, Fisher, & Duran, 1999).

The Functional Independence Measure (FIM™) Uniform Data System for Medical Rehabilitation (USD_{MR}), 1997) has also been used in determining amount of assistance needed. Raw FIM™ total scores have correlated well with estimated hours of care required ($r = -0.76$), however the association with estimated hours of supervision was poor ($r = -0.39$) (Disler, Roy, & Smith, 1993). Other studies have documented the association between the FIM™ total scores and minutes of assistance provided and global functional level (Deutsch, Braun, & Granger, 1996; Domerick, Edwards, & Diringer, 2003; Granger, Cotter, Hamilton, & Fiedler, 1990; Granger, Cotter, Hamilton, Fiedler, & Hens, 1993; Granger, Divan, Fiedler, 1995). In studies with persons with multiple sclerosis, stroke, and brain injury, it was found that a 1 point increase in the FIM™ total score was equivalent to approximately 2.19 to 5.12 minutes of care needed per day (Granger et al., 1990, 1993, 1995). Additionally, for a sample of persons within a rehabilitation center following cerebral vascular accidents, a 10-point improvement in the FIM™ total score was associated with a 50% decrease in the amount of care required (Dromerick, Edwards, & Diringer, 2003).

While the Functional Autonomy Measurement System (SMAF) was not specifically designed to predict level of dependence, an initial study documented that

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

there was a significant correlation between SMAF disability index scores and amount of nursing care required ($r = 0.88$, $p < .001$) (Hébert, Carrier, & Bilodeau, 1988). More recently, the SMAF was used to create 14 different disability profiles for older adults (Dubuc, Hébert, Desrosiers, Buteau, & Trottier, 2005). The characteristics of the individuals in the 14 different profiles range from needing supervision for housekeeping and difficulty with meal preparation, transportation, and budgeting (Iso-SMAF profile 1, mean score of 9.33), to being bedridden and dependent in ADL (Iso-SMAF profile 14, mean SMAF score of 73.77).

In addition, the Index of Independence in Activities of Daily Living (Index of ADL) (Katz, Ford, Moskowitz, Jackson, & Jaffe, 1963) levels have been concurrently correlated with amount of personal assistance required. That is, a grade of A correlated with independent functioning within the community and grades of E, F, or G correlated with needing nonfamily attendant care and a higher rate of nursing home placement (Katz, Downs, Cash, & Grotz, 1970). Lastly, the scores from the Klein-Bell ADL scale have been correlated with the hours per week of assistance one receives with ADL tasks ($r = -0.86$, $p < .001$); the lower the score on the Klein-Bell Scale, the greater the amount of assistance needed (Klein & Bell, 1982).

Such initial research has verified that assessments of ADL correlate with amount of care needed and/or community independence. However, additional research needs to be done to further verify their validity in specifically predicting the level of support that is needed for community dwelling. For example, the AMPS has been shown to predict the need for assistance, however predicting the amount and type of assistance needed has not been investigated. The FIM™ scores and scores on the Klein-Bell ADL scale

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

correlate with time of assistance required, yet this does not give information about the type or extent of assistance required. Lastly, the scores of the SMAF have been shown to significantly correlate with amount of nursing care, and more importantly, have been used to categorize level of disability in elderly long-term care residents. However, researchers have not investigated the use of the SMAF in categorizing level of disability in a more heterogeneous sample (i.e., those less than 65 years of age living in a long-term care facility).

It is important to note that no assessment of ADL would likely be a perfect predictor of the amount of assistance needed for community living, as tests of ADL have been designed to evaluate PADL and/or IADL ability, not level of dependence. Thus, it is important to stress that decisions about a client's need for assistance to live in the community are rarely made based on the results of one assessment. The occupational therapist often assesses the client's performance in a few ADL areas, gathers other pertinent information with regard to client's physical and social environments, formally or informally assesses social interaction skills, and gathers information from other healthcare providers to make decisions about the client's ability to perform all the daily life tasks necessary for independent living. Regardless, given the fact that occupational therapists often utilize assessments of ADL to determine a client's level of dependence, there remains a need to generate evidence of the extent to which the results of such assessments can *contribute* to the overall prediction of how much assistance the client will likely need to live in the community.

In review, informal investigation and review of the literature has revealed that (a) occupational therapists utilize assessments of ADL as a primary means of predicting

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

global functional level, (b) numerous ADL assessments significantly correlate with the need for assistance to live in the community, and (c) ADL assessments have not been shown to specifically predict the extent and/or amount of assistance individuals need to live in the community. This information has led me to the conclusion that there is a need to formally investigate if an ADL assessment can be used to accurately predict *how* much assistance a client likely needs to live in the community.

As it would be too cumbersome to investigate the concurrent and/or predictive validity of multiple ADL assessments, I have chosen to focus my research on the utilization of the AMPS in determining the amount of assistance that a client likely needs to live in the community. Although further explained in Chapter 2, I have chosen to investigate the AMPS because the AMPS offers a wide assortment of benefits including that (a) modern test theory has been utilized in the development and standardization process; (b) linear ADL ability measures are generated, and thus direct comparisons between individuals can be made (regardless of which tasks are performed); (c) the AMPS ability measures have been shown to be sensitive outcome measures; (d) an established cutoff measure has been developed that can be used to identify clients who likely need assistance to live in the community; (e) the AMPS can be utilized to evaluate nearly any person 3 years of age or older, regardless of diagnosis; (f) the AMPS is a performance-based assessment, not a questionnaire about one's performance; and (g) the AMPS is a client-centered, occupation-based analysis of the quality of ADL task performance (further clarification and discussion of the above points are included in Chapter 2). The AMPS ADL motor and ADL process ability measures were thus

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

investigated to determine if they can be validly used to assist in determining and/or predicting level of dependence.

More specifically, based the clinical use of the AMPS as an indicator the need for assistance, and the gaps in the available literature pertaining to determining level of dependence, the research questions in this investigation included: (a) Is there a significant association between ADL ability (motor and/or process) and global functional level?; (b) Can ADL motor and/or ADL process ability be used to correctly categorize individuals who need assistance and those who do not need assistance to live in the community?; (c) Can the ADL motor and/or ADL process ability measures be used to accurately determine level of dependence within the community (i.e., independent vs. minimal assist vs. moderate to maximal assistance)?; and (d) When using ADL motor and/or ADL process ability measures to determine level of community dependence, are there global differences in relative predictive validity across the different diagnostic groups? Through answering the above questions, I will provide occupational therapists with evidence of the validity and reliability of utilizing AMPS ability measures to support their clinical decisions regarding level of community dependence. If the AMPS ability measures show merit in predicting level of community dependence, the results of this research study will allow occupational therapists to more confidently "stand behind" their recommendations regarding their clients' need for assistance to live within the community. If the evidence does not support the use of the AMPS ability measures in predicting level of community dependence, then the results will lead researchers and occupational therapists to search out new strategies and/or new assessments that can

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

validly and reliably determine how much assistance their clients need to live in the community.

In order to investigate the predictive validity of AMPS ADL motor and ADL process ability measures, it was necessary to find a gold standard assessment that could be used as an external criterion of the amount of assistance needed for community dwelling. Obviously, the most ideal and accurate assessment of one's ability to live within the community would include direct observation of the client's ability to perform ***all*** of the necessary tasks and activities necessary for independent living (Dever & Knapczyk, 1988); however, this is not clinically practical. Upon investigating numerous assessments of independent living (see Chapter 2), none were deemed capable of accurately assessing *all* of the needed independent living skills. It appears that the existence of a reliable, valid, sensitive, and comprehensive assessment of one's ability to live independently within the community is either nonexistent or inaccessible via a literature search using of some of the popular databases (e.g., Web of Science, Medline). This discovery is in line with other researchers' findings (Dickerson, 1997; Iyer, Rothmann, Vogler, & Spaulding, 2005; Townsend & Ryan, 1991), for within their investigations, they also failed to find an adequate assessment of global function within the community.

Alternative options to utilizing a gold standard assessment of global function within the community include using the client's current living environment, self-report, proxy-report, and/ clinical judgment. Some have documented that there are inconsistent levels of dependence between and within seemingly similar residential settings. For example, Willer and Gustaferro (1989) stated that it is possible that "the label used to

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

identify a setting does not necessarily indicate the level of independence experienced by the individual living in that setting" (p. 273). Given that the current living environment does not necessarily depict the amount of assistance that an individual needs to live in the community, there appears to be little value in utilizing this as a gold standard assessment of level of dependence.

In examining the use of proxy (caregiver) report, one inherent problem is that many individuals, although in need of some assistance, may live alone and thus may not have an available proxy to report their level of dependence. Additionally, caregivers may have a tendency to declare higher rates of dependency (Cotter, Burgio, Stevens, Roth, & Gitlin, 2002; Santos-Eggimann, Zobel, & Bérod, 1998; Rubenstein, Schairer, Wieland, & Kane, 1984). In contrast, when considering the accuracy of self-report questionnaires, there may be a tendency for clients to underestimate levels of dependency and/or disabilty (Cotter et al., 2002; Poole, Atanasoff, Pelsor, & Sibbitt, 2006; Santos-Eggimann et al., 1998; Rubenstein et al., 1984). Thus, it appears that the utilization proxy and/or self-report questionnaires may not offer the most accurate rating of level of dependence.

The utilization of clinical judgment of occupational therapists has shown more promise in accurately assessing level of community independence, provided the therapist focuses on occupational performance rather than body function(s) (Pinholt, Kroenke, Hanley, Kussman, Twyman, & Carpenter, 1987; Rogers et al., 2003). For example, in a study that examined the accuracy of medical doctors in rating global functional status, it was noted that the doctors were able to identify the need for assistance in most areas of basic ADL (feeding, bathing, dressing, and toileting) with sensitivity ranging from 81% to 92% (Pinholt et al., 1987). It stands to reason that if medical doctors are able to

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

identify the need for assistance in ADL (within a reasonable range of error), that occupational therapists, who are specialists in the assessment of ADL ability, would likely be even better judges of global functional status/level of dependency.

While there has been some criticism in using clinical judgment to generate outcome measures (Grove & Meehl, 1996), there are several examples of the use of clinical judgment as an external criterion for a previously standardized tool (Donkervoort, Dekker, Stehmann-Saris, & Deelman, 2001; Forette, Panisset, & Boller, 1992). Additionally, others have documented high inter-rater reliability for expert ratings of global function (Dromerick et al., 2003; Whelihan, Emerson, Kleban, & Granick, 1984). Lastly, in their study of clients with dementia, Hartman et al. (1999) used occupational therapists' clinical judgment via utilizing the AMPS global ratings of functional level, documenting that the potential for bias was minimized due to the fact that at the time of documenting the client's global functional level, the clinician was blind to the client's ADL motor and ADL process ability measures.

As clinical judgment has shown some promise in determining global function within the community, I concluded that this seemed to be one of the most ecologically valid and accessible means of identifying how much assistance individuals need to live in the community. With this in mind, I further investigated the AMPS global functional level ratings (clinical judgments) to see if this rating system could be used in the current research project as an external criterion of level of dependence.

The AMPS global functional levels are used to rate the client as either (a) able to live independently in the community, (b) needing minimal assistance to live in the community, or (c) needing moderate to maximal assistance to live in the community

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

(Fisher, 2006a, 2006b) (Table 1). When determining the AMPS global functional level rating, the therapist must consider the client's ability to perform ***all*** of the necessary tasks and activities that are required for independent living (Fisher, 2006a, 2006b). In fact, the specific instructions in the AMPS manual state that the judgment of the client's over all functional level "should be based on all information available to the examiner and should not be based solely on the client's AMPS task performance" (Fisher 2006a, p. 86). As a result, the therapist must utilize multiple sources of information to determine the most accurate rating. For example, the clinician may utilize information gathered through (a) interviewing the client and/or caregivers, (b) interviewing other healthcare providers (e.g., case managers, nurses, physical therapists, speech therapists, doctors), (c) observing and assessing the client's actual ADL ability, and/or (d) assessing available resources as well as the client's body functions. This approach is in line with current trends, for it is thought that the utilization of multiple sources may be the most accurate means of determining level of dependence and global function (Dickerson, 1997; Iyer et al., 2005).

One may question whether the AMPS global functional level ratings are biased by the AMPS evaluation (or visa versa). First, it is important to note that at the time of rating the client's global functional level, the AMPS raters were (a) blind to the person's final AMPS ability measures, and (b) not aware that the functional level ratings would be used to investigate the validity of the AMPS ability measures. Lastly, it is important to note that the investigation of the validity of the AMPS global functional level ratings leads to a "catch 22" situation, in that a gold standard with which to compare the accuracy of this rating system has not been found.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

As stated above, the AMPS rater is blind to the client's ability measure. It is important to realize that the AMPS computer-scoring software (Three Star Press, 2005) utilizes a complex mathematical calculation to adjust the client's ability measure based on (a) the challenge of the tasks performed, (b) the severity of the rater, and (c) the difficulty of the ADL motor and ADL process skill items. Thus, it is virtually impossible for an AMPS rater to predict the client's linear ability measures based on the raw item or total scores. Furthermore, one must understand that the AMPS global functional level rating is based on the clinician's judgment of the client's ability to live within the community, and such judgments are derived from collecting as much information as possible, from multiple sources, about the client and his/her specific abilities and environmental constraints.

Finally, there is no indication within the AMPS manual or the professional literature regarding where individuals of varying functional levels fall on the ADL motor and/or ADL process ability scales. The AMPS Graphic Report (Appendix B) is currently used to illustrate more or less ADL motor and ADL process ability, and if a client falls below the ADL motor and/or ADL process cutoff measure, there is some indication that the client likely needs some assistance to live within the community (Fisher, 2006a). However, there is no indication of *how much* assistance the client needs, based on his or her ADL ability measure(s). As a result, the clinicians were naïve to how the global functional level ratings would impact the determination of specific levels of community independence/dependence along the continua of linear ADL motor and ADL process ability measures. Thus, within this study, the most practical and valid option available

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

was to utilize the occupational therapists' clinical judgment of global function through the utilization of the AMPS global functional level ratings.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

CHAPTER 2

LITERATURE REVIEW

Independent Living

As one considers the concept of assessing independent living skills, one must first come to understand what is meant by the term "independent living." According to Friedan, Bachelor, Cole, and Bailey (as cited in Bachelder, 1985), independent living is defined as:

> Control over one's life based on the choice of acceptable options that minimize reliance on others in making decisions and in performing everyday activities [tasks]. This includes managing one's affairs, participating in day-to-day life in the community, fulfilling a range of social roles, and in making decisions that lead to self-determination and the minimization of physical and psychological dependence on others. (p.100)

As one can see, the definition of independent living is very broad and complex. Thus, the determination of the extent to which an individual can attain these goals is quite complicated. In order to assess one's ability to live independently, one must acquire an understanding of the necessary tasks, that when performed in an effective manner, constitute successful community and/or independent living. In order to be thorough and critical within this investigation, it was important to first determine the constellation of tasks that are included within the realm of independent living.

Many authors have similar beliefs regarding the basic tasks that one must perform in a competent manner in order to live independently. For example, several authors have identified the following areas as being vital for independent living: (a) basic self-care (e.g., dressing, grooming, hygiene), (b) food preparation, (c) home maintenance (e.g.,

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

cleaning the house, caring for clothes, maintaining the yard), (d) transportation and community mobility, (e) money management, (f) leisure, (g) health and safety (e.g., first aid, emergency procedures, knowledge of proper nutrition, drug/alcohol awareness, safe sex), (h) shopping, (i) social interaction, and (j) self-direction (Agran, et al., 1994; DeLoach et al., 1983; Dever & Knapczyk, 1988; Globe Fearon, 1997; Webber et al., 2002).

Although there is basic agreement among authors, Dever and Knapczyk (1988) have documented one of the most systematic and thorough taxonomies of necessary community living skills. Dever and Knapczyk's taxonomy of community living skills includes the following five domains (a) personal maintenance and development, (b) homemaking and community life, (c) leisure, (d) travel, and (e) vocational.

Although Dever and Knapczyk's (1988) taxonomy includes a vocational domain, within the realm of the current project, employment was not considered to be a necessary activity for community living. While employment constitutes a very important and meaningful aspect of one's life, it is not a prerequisite for living independently in the community. For example, many older adults who have retired, or have chosen not to work, continue to live independently in the community, as do many younger individuals (e.g., homemakers, individuals with disabilities). Moreover, literacy and prevocational skills do not seem to correlate with independent living skills (Evenson & Boyd, 1993). Thus, the first four of Dever and Knapczyk's domains will be the focus of the remainder of this discussion.

The domain of personal maintenance and development includes caring for one's health and personal activities of daily living (PADL) (e.g., bathing, grooming, toileting)

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

and interacting with friends and family members (Table 2). The domain of homemaking and community life includes obtaining and maintaining one's living quarters, home safety, meal preparation, money management, and social interaction (Table 3). The leisure domain (Table 4) involves selecting and performing leisure activities in an effective and appropriate manner, and lastly the travel domain involves traveling to and within buildings and community settings (Table 5). In the most general sense, the specific items within each domain revolve around the necessary tasks that must be performed, dealing with and preventing accidents/emergencies, dealing with unexpected occurrences (e.g., equipment breakdowns, disrupted routines), and interacting appropriately with others.

The taxonomy created by Dever and Knapczyk (1988) is organized in a manner that represents the complex and interactive nature of the five different domains. Although, the five domains can be operationally defined as separate entities, the domains exhibit a certain degree of overlap. For example, as one partakes in a leisure activity within the community, he/she must dress in reasonably clean clothes that are appropriate given the weather conditions and activity choice, and he/she must travel to and within the leisure site. Thus, in this instance, the leisure domain overlaps with the domains of personal maintenance and development, homemaking and community life, and travel.

Although Dever and Knapczk (1988) have provided the most thorough taxonomy of community living skills found during an extensive search of the literature, the taxonomy may have a few shortcomings. For example, within the domain of personal maintenance and development (Table 2), the task of dressing one's self is not considered. Within the their personal maintenance domain, one only considers appropriate dress, not

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

the act of dressing. Within the realm of independent living, however, one must not only choose appropriate clothing, one must also effectively don/doff clothing to be considered independent within the community. Another potential shortcoming is that some of the items listed in Tables 2 through 5 may be out-dated and/or culturally bound, and thus one may need to investigate the need to revise some of the items so that they are in line with current societal and cultural trends.

Through investigating the multifaceted and complex nature of independent living, one can recognize that the skills involved in maintaining a successful life within the community are very intricate. Given that there are many different and diverse skills that one must acquire and develop in order to live independently, the act of accurately predicting ability to live independently in the community becomes a tremendous undertaking. However, with an understanding of the concept of independent living, one can begin to evaluate and critique the available assessments that are currently being used within the healthcare community to evaluate and predict level of independence in the community.

Modern Measurement and Test Theory

Within the following sections, I will explore and critically evaluate a number of activities of daily living (ADL) and independent living assessments. As I do this, it is important to understand that I come from a background that endorses the philosophical and mathematical concepts of modern test theory and objective measurement. Throughout my evaluation of the assorted assessments, the concepts and beliefs regarding objective measurement often provide a foundation and rationale for my critique. Thus, prior to delving into my evaluation of the assessments, it is important to first provide the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

reader with the background information needed to understand what is meant by objective measurement. Within my discussion of objective measurement, I will start with fairly broad philosophical explanations and then I will carry these notions forward into a more specific explanation of objective measurement and the basic assumptions of the family of Rasch models.

One of the most important tasks that researchers face is conveying their discoveries to others in the most accurate and truthful manner possible. Those who utilize qualitative methods as a means of investigation likely convey their findings in written text, while those who utilize quantitative methods tend to convey their discoveries through the use of written and numerical symbols. Regardless of how ideas are conveyed, researchers must strive for clarity; they must strive to use symbols (e.g., words, numbers, pictures) that have generally agreed upon meanings. If the methods of exploration and the dissemination of the results involve the utilization of ambiguous symbols, one must recognize that multiple meanings can be generated (Fisher, 2002).

During the course of history, a detrimental error occurred within the minds of researchers and the population at large. That is, it became common practice to infer that numbers were innately objective and accurate in conveying meaning. To understand this error, and the overarching concept of objective measurement, one must first come to realize that meaning is an internal process that is generated within the mind of the beholder. Thus, symbols do not directly generate meaning; people with unique historical perspectives generate meaning when they see, hear, and/or read symbolic configurations (Fisher, 2002). Thus, just as multiple meanings can be generated by individuals

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

partaking in a conversation, reading a book, or perhaps viewing a piece of artwork, multiple meanings can also be generated from numeric symbols.

As a means of fully understanding what is meant by objective measurement, one must not only acknowledge the notion of symbol–meaning independence, one must also recognize that in order to generate meaning, we must have prior knowledge about the symbol. When a person is presented with a situation or symbol that is unfamiliar, he/she must employ that which he/she already knows in order to understand and comprehend the information. It is virtually impossible to interpret a new situation or a symbol any other way, for we can only start with what we already know. For example, if I am presented with an unfamiliar word, I will try to figure out what the word means by looking at the familiar words within the sentence and by examining the context of the sentence. I may then turn to the dictionary to seek a definition. When I read the definition of the word, I am in a sense striving to use my existing knowledge and terminology to better understand the meaning of the unfamiliar word. If the definition is riddled with even more unfamiliar and/or ambiguous words and terminology, then I will not be able to grasp the concept of the new word. If I am able to understand the definition, because my existing knowledge and historical background support this, then I will gain new insights and generate my own personal meaning of that word. If there is strong symbol–meaning coordination, then the meaning that I generate will be closely linked to that of my peers, and thus, we will all share a common meaning when we read/utilize that specific word (Fisher, 2002).

Likewise, if I were to come across a new or unfamiliar numerical measure, I would need to utilize my prior knowledge to gain an understanding of the meaning

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

behind the number. For example, many countries outside of the United States use the metric system. When I am outside the United States, I find myself constantly using my current knowledge to understand the meaning of common, but unfamiliar measures. For example, if I am in Sweden and I note that the current temperature is 8 degrees Celsius, I must convert the measure into Fahrenheit in order to gain an understanding of what this temperature means. That is, I have to convert the unfamiliar measure into one that is more compatible with my current knowledge. If, however, I were to repeatedly utilize Celsius as a measure of the temperature, over time I would no longer need to convert the measure, for I would begin to develop a strong symbol–meaning coordination of. Thus, in the later instance, my current knowledge would have expanded to incorporate new concepts and new measures, such that both measures of temperature (Fahrenheit and Celsius) would have strong symbol–meaning coordination.

If we continue to build on this line of reasoning, we must recognize that for the ancient Greeks, the term mathematics signified those things that are already known (Fisher, 2002). Numbers became strongly associated with mathematics because they were the most transparent symbols (Fisher, 2002). Numeric symbols offered scientists and researchers a means to achieve rigorous symbol–meaning coordination, for researchers were able to clearly state the amounts and quantities of things of interest. Perhaps the pervasive symbol–meaning coordination and the use of numbers within the realm of physical sciences lead social scientists astray in that they developed a false belief that numbers inherently convey objective meaning (Bond & Fox, 2001; Fisher, 2002; Michell, 1999, 2000). However, over time, the numerical representation of quantity within the human sciences became increasingly ambiguous and meaningless, for

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

researchers failed to demonstrate strong/valid symbol–meaning coordination through the use of numerical signs (Fisher, 2002; Michell, 1999, 2000).

If we dig a bit deeper into this issue, we can begin to realize that "quantitative" research is not always inherently mathematical in nature, for when numerical signs have weak symbol–meaning coordination, multiple meanings can arise from the data. For example, suppose a researcher analyzed a group of five clients to determine their ADL ability. If she were to simply rank these clients along an ordinal scale from most able (5) to least able (1), there is no way of determining how much better (or worse) one client is compared to the others. Thus, there is a great deal of room for interpretation within these data. For example, the client ranked as number 5 may be much more able than the 4th ranked person, yet the 4th ranked person may have very similar abilities to the 3rd ranked person. As a result, the differences between rankings of 5 and 4 are not equal to the differences that exist between the rankings of 4 and 3. It stands to reason, therefore, that the use of a numerical rank order is a non-mathematical, weakly coordinated symbol–meaning system (Bond & Fox, 2001; Smith & Smith, 2004; Wright & Masters, 1982; Wright & Stone, 1979).

A truly mathematical approach to this problem would be to devise a linear, additive scale of ADL ability, for within this system, the numerical signs would clearly indicate exact differences in ability. With the use of a linear scale, one would have direct knowledge about how much more able one person is in comparison to another (Bond & Fox, 2001; Smith & Smith, 2004; Wright & Masters, 1982; Wright & Stone, 1979). In this mathematical system, there is strong symbol–meaning coordination, for one enters into the problem with a sense of what the numbers truly represent in terms of ADL

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

ability. Furthermore, there is considerable agreement within the literature that if one intends to use measures to infer, that the measures must be equal interval (Bond & Fox; Smith & Smith; Wright & Masters; Wright & Stone).

As a means of delving into the realm of objective measurement, one must first define what is meant by objectivity. According to Fisher (2000), "objectivity is characterized by the separation of meaning from the geometric, metaphoric, or numeric figure carrying it, allowing an ideal and abstract identity to take on a life of its own" (p. 527). Bond and Fox (2001) state that objective measurement "requires that the measure assigned to the attribute/construct be independent of the observer" (p. 2). Thus, if a symbol is used to signify the extent or amount of a given trait, then the meaning of that symbol must be independent from the person being evaluated (object of interest) and from the person measuring the trait. Only through such objectivity can we generate a shared meaning of what it means to have a certain amount of a given trait. If the objective measure (symbol) is not independent of the object of interest, there is room for interpretation and the generation of multiple meanings.

As this is a complex concept, it may be easier to explain through the use of a practical example. If we state that Jack is the tallest boy in his class, this is not an objective measurement, for this statement does not separate Jack's apparent height from Jack and Jack's environmental conditions, for rank order is always relative to the observer and the observed. In this instance, we have described Jack in relation to his peers, yet we do not know how tall he is, nor do we now the height of his peers. If we do not know anything about Jack or his peers, we have no idea of how tall he is; we are simply left to ponder this question and develop our own ideas about his height. Likewise,

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

if John says, "Jack is really tall," we again do not have a clear definition of Jack's height, as this concept could vary depending on John's beliefs and/or his own personal traits. For John, who might be very short, may consider anyone taller than himself to be "very tall." If, on the other hand, we state that Jack is 5 feet 11 inches tall, and we also have a record of how tall his peers are (5 feet 3 inches to 5 feet 10 inches), then we can develop a common idea as to how much taller Jack is in relation to his peers. The measurement of 5 feet 11 inches is independent of Jack and independent of the person measuring his height.

Within an American culture, most individuals have a fairly good idea as to how tall one is if he is 5 feet 11 inches, as there is a common understanding of length measured in feet and inches. Thus, when we hear that Jack is 5 feet 11 inches tall, we can converge on a similar understanding of Jack's height. The point to stress is that since we cannot separate Jack, the individual, from the "tallest boy in the class" (for they are one in the same), we cannot develop a shared meaning of his height from the statement "Jack is the tallest boy in his class", and thus it is not an objective measurement.

If we consider the statement by Bond and Fox (2001), that measures "must be objective abstractions of equal units" (p. 2), we can begin to develop an understanding of what this statement truly means. First and foremost, as stated above, if we utilize numerical symbols to convey meaning, the most coordinated symbol–meaning system consists of linear, additive (interval) scales (Bond & Fox, 2001; Fisher, 2002; Smith & Smith, 2004; Wright & Masters, 1982; Wright & Stone, 1979). When measures are plotted along a linear continuum (interval scale), one can easily determine how much

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

more or less of the attribute one possesses. Stated differently, there is greater symbol–meaning coordination when one utilizes a linear scale.

Secondly, when clearly defining objective measurement, one must also come to understand what the authors mean when they refer to "objective abstractions" (Bond & Fox, 2001). To best explain this concept, consider the following example. If I asked a peer to "bring me 5 feet 11 inches," she would look at me as if I were crazy, for this task is virtually impossible. Sure, she could bring me a piece of rope that is 5 feet 11 inches long, but this is not the measurement itself, it is simply a piece of rope. Within our minds, the concept of 5 feet 11 inches is real, yet it is merely an abstract concept that has common meaning and understanding within one's culture.

The same phenomena can be found in language. That is, I could also ask my peer to bring me "green." She could of course bring me a handful of grass, or she could bring me a green crayon, but she could not bring me "green," for green only exists within our minds. I could objectively state that the grass is green, and very few people would argue this point, for we have a shared belief system as to the nature of greenness. Thus, when we achieve a strong shared belief about the meaning of words or measures, and when there is robust symbol–meaning coordination, we have achieved what is referred to as "objective abstractions."

Fisher (2000) pointed out that there are at least two aspects of objectivity that must be considered, that of social objectivity and that of objective methodology. If we first consider social objectivity, we realize that symbol–meaning coordination is a social construct that must be developed through the work of metrology (Fisher 2000, 2002). Metrology is a term used to describe the purposeful development of symbol–meaning

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

coordination. In other words, it is the development of a "common language" regarding the meaning of measures. In order to accomplish this, scientists and researchers must make considerable efforts to communicate and develop systems of understanding. Instruments must be evaluated within and between laboratories, among different examiners, and with different samples to determine if it is valid and reliable. If the instrument proves its worth, then scientists must also make efforts to promote the instrument and the resulting measure, for only through a coordinated effort between researchers and scientists can a measure become objective (i.e., have strong symbol–meaning coordination). If a scientist works in isolation developing an instrument for measuring a specific trait, there is little to no chance that he will be able to develop a truly objective measure. That is, the symbol–meaning coordination would not exist outside the bounds of his laboratory, and there would be little proof that his device withstands the test of invariance across samples, environmental conditions, and examiners.

The second aspect of objectivity is that of objective methodology (Fisher, 2000). Within this concept, Wright and Stone (1979) stated that an objective measure must be one in which the calibration of test items is independent of the sample (i.e., sample-free item calibration), and one in which the measurement of the object (or person) is independent of the measuring instrument (i.e., test-free person measurement). Thus, in order to generate an objective measure, individuals must respond in a predictable manner, regardless of personal characteristics, such that the relative difficulty of the items is similar across all persons (Wright & Stone, 1979). In fact, Wright and Stone stated that "calibration of test-item difficulty must be independent of the particular persons used for

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

calibration . . . . This is the only way we will ever be able to construct tests which have uniform meaning regardless of whom we choose to measure with them" (p. xii).

As sample-free item calibration and test-free person measurement are integral to Rasch measurement models, this seems like a good place to introduce the idea that researchers can utilize the family of Rasch models to obtain objective measures. One large advantage of utilizing a Rasch model is that researchers have the ability to convert ordinal data into linear ability measures. Thus, through Rasch analysis, the potential for generating objective measures is greatly enhanced.

As one enters into the task of developing an objective assessment tool or instrument, great care must be taken to first develop a theoretical model of the construct to be measured (Bernspång & Fisher 1995b; Fisher, 1993; Fisher, 2002; Wright & Stone, 1979). One must conceptualize a line or continuum of the given trait or ability, and determine what it would look like to have more or less of this trait. For example, if we set out to measure one's ability to ride a bicycle, we can quickly start to develop a notion that one with greater cycling ability will be able to ride longer distances, climb steeper and longer hills, safely ride on uneven terrain, generate higher speeds for longer periods of time, and maneuver through corners and obstacles with ease. When envisioning a person with the most amount of cycling ability, one may think about the traits and abilities of someone like Lance Armstrong (seven time consecutive Tour de France winner). When envisioning a person with very little cycling ability, one may consider a child who is just learning how to ride without training wheels, who can only travel a few feet before he falls over or before his parent must hold onto the bike to prevent him from falling.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

As one can see, there is theoretically a large continuum of cycling ability ranging from a child just learning how to ride a two-wheeled bike all the way up to someone with the cycling skills like Lance Armstrong. Within the conceptualization of this continuum of cycling ability, one could also begin to develop a continuum of items/tasks that range from easy to hard. For example, riding one block on a two-wheeled bike should be easier for every person than riding 100 miles of varying terrain, and riding on a flat surface should be easier for all persons than cycling up a steep hill. Likewise, riding at a speed of 5 miles per hour on a flat road for 10 minutes should be easier for all persons than riding at a speed of 20 miles per hour for 10 minutes on a flat road. Those with greater cycling ability are more likely to pass (or endorse) all items, and all persons (regardless of ability) are more likely to pass or endorse the easy items than they are to pass the hard items. In theory, this is what one would expect to happen if we set out to measure cycling ability.

Within this simple example, I have begun to illustrate the basic assertions of a many-faceted Rasch model which include: (a) all persons are more likely to receive higher scores on easier items/tasks than harder items/tasks, (b) persons of higher ability are more likely to receive higher scores on all items/tasks than are persons of lower ability, and (c) strict raters are more likely to assign lower scores across all items/tasks than are lenient raters. The last assertion was not illustrated in the above example, as it only applies to many-faceted Rasch models where rater severity is taken into consideration.

Once we have created the assessment, we would need to test a group of cyclists of varying ability. Upon entering the data into a Rasch-specific data analysis computer program, such as FACETS (Linacre, 2005), we could analyze the output to see if our

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

theoretical model holds true. For example, we would want to investigate if the items calibrated in a predictable manner along a continuum from easy to hard. We would also want to investigate the items to see if there was a predictable response pattern. For example, did all individuals consistently obtain higher scores on easier items? Did cyclists that are more able obtain higher scores across all items than did less able cyclists? We would also want to examine the resulting ability measures of the cyclists to see if they were located on the continuum where we would expect. For example, we would expect a 4-year-old child to have less cycling ability than a teenager or adult. Likewise, we would expect a competitive cyclist to have more ability than a casual rider. Overall, the goal is to determine if there are any violations of our theoretical model. If the basic assertions have not been violated, we can be more certain that the assessment has produced objective measures, for these measures would be based on sample-free item calibration and test-free person measurement (Wright & Stone, 1979).

Furthermore, during the development of an assessment tool, one must strive to create items that only contribute to the construct of interest (e.g., cycling ability), so that a unidimensional scale of that construct can be generated. For example, if we believed that vegetarian cyclists possess more cycling ability, we could include an item on our assessment that asked about the eating habits of cyclists. When conceptualizing this item, we would predict that it would be harder to endorse, in that less able cyclists would be less likely receive higher scores on this item than more able cyclists (e.g., less able cyclists would not maintain a vegetarian diet). If, however, upon further investigation, we discovered that within the sample of more skilled cyclists, some received high scores and some received low scores, and among the less skilled cyclists, there is a similar

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

pattern of both high and low scores, we would conclude that this item does not hold true to our model. Thus, vegetarianism does not contribute to our unidimensional scale of cycling ability. This item would be eliminated from the assessment, for within the family of Rasch measurement models, this item would be considered a "misfit" as noted by the item's mean square and standardized z values (Wright & Stone, 1979; Wright and Masters, 1982). The mean square and standardized z are goodness-of-fit statistics that help researchers evaluate the existence of a unidimensional scale.

Through objective methodology, we can determine and verify that our theoretical model holds true for a variety of persons, within different contexts, and between a diverse population of examiners; and that our tool appears to accurately measure cycling ability. The next step would be to promote the use of this assessment, such that within the cycling community there was sound symbol–meaning coordination with regard to the linear measure of cycling ability. For example, if it was widely known that Lance Armstrong has a cycling ability of 4.5 cycling units, a local male professional bike racer typically has an ability measure of about 3.0, a recreational cyclist typically has an ability measure of about zero, a sedentary adult with little cycling experience has an ability measure of about -1.5, and a child just learning to cycle typically has an ability measure of about - 4.0, etc., then it would be relatively easy to interpret measurement values of those who were subsequently tested. Continuing with this example, individuals within the cycling community would have a common language for documenting and discussing cycling ability. Thus, instead of stating "he is a really good cyclist," one could state his cycling ability measure is 2.5, which would conjure up a common image of a fairly skilled and talented cyclist.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Through creating a common language and strong symbol–meaning coordination for the linear scale of cycling ability, those within the cycling community could achieve social objectivity. In summary, through attending to both objective methodology (e.g., sample-free item calibration and test-free person measurement) and social objectivity (e.g., developing a common language and strong symbol–meaning coordination), we theoretically could develop a truly objective measure of cycling ability.

Within the realm of social sciences, further efforts need to be taken so that we can continue to create truly objective measures. Within the field of occupational therapy, the utilization of objective measures has begun to raise the quality and accuracy of our assessments and our predictions; however, continued efforts need to be made to ensure that researchers are utilizing and promoting objective methodology and social objectivity. Objective measures can and should provide therapists with a common language and a sense of shared meaning regarding the measures that their clients attain. In my opinion, this could, in part, help to give occupational therapists greater power in describing their unique contribution within the healthcare arena. With the concept of objective measurement in mind, my next step will be to critique and summarize the means by which occupational therapists assess, measure, and/or predict level of dependence with regard to community living.

Although my research questions focus on the utilization of the Assessment of Motor and Process Skills (AMPS) (Fisher, 2006a) in predicting level of dependence, I felt that it was necessary to also review other ADL assessments to provide a foundation of comparison. Thus, in the following section I have included an extensive review of the AMPS as well as a review/critique of other ADL assessments. Although not all of the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

available ADL assessments used by occupational therapists are discussed, the following provides a representative sample of the types of assessments that are used.

Assessments of Activities of Daily Living

Assessment of Motor and Process Skills (AMPS)

The Assessment of Motor and Process Skills (AMPS) is a standardized occupational therapy assessment that is used to gain information about the quality of an individual's performance of chosen, familiar ADL tasks (Fisher, 2006a, 2006b). With the use of the AMPS, occupational therapists document the quality of ADL performance by rating the amount of effort, efficiency, safety, and independence that the person displays during the performance of the 16 ADL motor and 20 ADL process skill items defined within the AMPS (Table 6).

More specifically, the AMPS is administered by trained and calibrated occupational therapy practitioners according to standardized procedures described in the AMPS manual (Fisher, 2006a, 2006b). An AMPS evaluation consists of several steps, the first being the occupational therapy interview. During this interview, the therapist inquires about the types of tasks that the client performs on a daily basis and those tasks that are presenting a challenge to the client. The therapist determines which personal and instrumental ADL tasks the client is interested in and familiar with, and the therapist chooses several (usually three to six) appropriately challenging tasks from among the 83 ADL tasks listed in the AMPS manual. From this shortened list, the client chooses at least two tasks to perform, and the therapist and the client agree upon the specific task criteria for each of these tasks. The client performs each chosen task in a familiar environment and in his or her usual manner. After completion of each of the AMPS

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

tasks, the client is scored on 16 ADL motor and 20 ADL process skill items according to the criteria in the AMPS manual (Fisher, 2006b) using a 4-point ordinal scale. A score of 1 indicates deficient performance and/or task breakdown, a score of 2 indicates ineffective performance, a score of 3 indicates questionable performance, and a score of 4 indicates competent task performance.

It is important to note that the AMPS is a criterion-referenced measure of ADL ability. Thus, individuals are scored according to an external criterion of competence. As a comparison, a norm-referenced assessment compares an individual's performance to his/her peers or normative group. The results from a normative test do not provide information about specific abilities or skills, they are merely used to illustrate the person's relative performance as compared to others. Developing a criterion-referenced assessment is accomplished via creating a test that is "deliberately constructed to yield measurements that are directly interpretable in terms of specified performance standards" (Berk, 1984, p. 12). Within the AMPS, the ADL motor and ADL process skill scores are not based on how a "typical" person performs; the scores are based on the criterion of no observable increased effort, decreased efficiency, safety risk, and/or need for assistance (i.e., competence) (Fisher, 2006a). Therefore, a score of 4 represents safe, independent, efficient, and effortless performance of the ADL motor or ADL process skill.

As a means of gaining a better understanding of what the AMPS does and does not assess, it is important to explain in detail what is meant by performance skills. The ADL motor and ADL process skills are analogous to the goal-directed actions defined under the Activities and Participation domains of the *International Classification of Functioning, Disability and Health* (World Health Organization [WHO], 2001) and the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

performance skills in the American Occupational Therapy Association (AOTA) practice framework, (AOTA, 2002; Fisher, 2006c).

To further explain this concept, I will first globally define ADL motor skill and ADL process skill and then I will use an example to describe a few of the discrete skills or actions that one can observe during a task performance. ADL motor skills are defined as "the observable, goal-directed actions that a person enacts during the performance of ADL tasks in order to move oneself or the task object" (Fisher, 2006a, p. 4). For example, within the context of preparing toast with strawberry jam and a cup of coffee, we can observe the person *walking* to the cupboard, *pulling* the cupboard door open, *reaching* for a coffee cup, *grasping* the cup, *lifting* the cup, and *transporting* the cup to the table. ADL process skills are defined as "the observable actions of performance the person enacts to logically sequence the actions of the ADL task performance over time, select and use appropriate tools and materials, and adapt his or her performance when problems are encountered" (Fisher, 2006a, p. 4). As a person is preparing toast and coffee, we can observe the person *searching* for, *locating*, and *choosing* the strawberry jam, *gathering* the jam to the workspace, *pausing* before opening the jar, *supporting* the jar of jam so that it does not slip, and *using* a spoon to scoop the jam onto the toast.

As illustrated in the examples above, we begin to see that the ADL motor and ADL process skills are the smallest observable units of performance, and that when they are compiled together, they result in the task being completed (Fisher, 2006a, 2006c). It is important to recognize that ADL motor and ADL process skills are discrete actions of performance that are observed as a person is engaged in meaningful daily life activity.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Thus, they cannot be equated with body functions (e.g., strength, range of motion, cognition, visual perception, attention) (Fisher, 2006a, 2006c).

Consider another example. A client comes into the occupational therapy clinic, and the client tells the therapist that she cannot reach the bowls in her cupboard. If the therapist were to directly enter into assessing upper extremity strength and range of motion, she may ask the client to raise her arm above her head, and hold her arm up against resistance. Contrary to her assumptions, the therapist may find that the client has no strength and/or range of motion limitations. Thus, after 10 minutes of assessing the client's upper extremity body functions, the therapist is no closer to determining the reasons for her client's inability to reach the bowls.

If, on the other hand, the therapist began her assessment by observing the client make a bowl of cereal, she would observe that the client does in fact need assistance to obtain a bowl from the cupboard. After recording and documenting the client's reaching/positioning skill (occupational performance) deficits, the occupational therapist interprets the potential cause. In this case, the therapist observed that the client was readily able reach out and grasp task objects, thus she reasoned that the client's skill deficits were due not due to musculoskeletal impairments. Rather, she reasoned that the client's skill deficits could be attributed to environmental factors, for her walker tray prevented her from getting close enough to the cupboard to be able to reach the shelf where the bowls were located. In summary, even though her upper extremity range of motion and strength were within functional limits, she still could not reach and obtain a bowl from the cupboard. Thus, within this task performance and within this environment, the client has a reaching skill deficit. In this scenario, the therapist

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

observed the client as she was engaged in the performance of a meaningful and relevant task, versus simply observing and documenting body functions. It is my hope that through such examples that the reader will begin to more fully understand that the AMPS cannot be considered a test of body functions, rather, it is used to assess performance skills at the level of participation (Fisher, 2006a).

Returning to the standardized administration of the AMPS, once the occupational therapist has scored all of the ADL motor and ADL process skill items, the raw scores for each task performed are entered into a specialized application of many-faceted Rasch (MFR) analysis (Linacre, 2005), the AMPS computer-scoring program (Three Star Press, 2005). This program converts the raw ordinal ADL skill item scores into linear ADL ability measures (Fisher, 1993, 1997). This analysis adjusts the final ADL ability estimates to account for task challenge, ADL skill item difficulty, and the severity of the rater.

The resulting ability measures are plotted on linear continua of ADL motor and ADL process ability on the AMPS Graphic Report (Appendix B). The ADL motor and ADL process ability measures have been shown to correlate significantly with global functional level ($r = .48$, $p < .01$ and $r = .59$, $p < .01$, respectively) (A. G. Fisher, personal communication, January 13, 2005). More specifically, if a client falls below the current ADL motor cutoff measure of 2.0 logits, he or she is likely experiencing increased effort and potentially decreased safety and/or independence within the performance of ADL tasks. Additionally, an ADL process ability measure below the current cutoff of 1.0 logit indicates that the client is likely experiencing decreased efficiency, safety, and/or independence during ADL task performances (Fisher, 2006a).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Furthermore, ADL process ability has been shown useful when determining if a client likely needs assistance to live within the community. In an earlier study, approximately 93% of those with measures below the 1.0 logit ADL process ability cutoff required assistance to live in the community (Fisher, 2006a). Other researchers have investigated the reliability of the AMPS cutoff measures in predicting the need for assistance in clients with Alzheimer's disease (Hartman et al., 1999) and by utilizing the ADL process cutoff measure of 1.0 logits, they were able to accurately classify 94% of their sample as either needing or not needing assistance to live in the community.

In addition to the current AMPS cutoff measures, researchers have defined a potential risk zone that is defined as ability measures that are within ±0.30 logit of the ADL motor and/or ADL process cutoff measures (Bernspång & Fisher 1995a; Hartman et al., 1999). Clients that have ability measures within the ADL motor risk zone may be at risk for functional decline, in that they are likely beginning to show signs of increased effort and possibly decreased safety during ADL task performance. Likewise, ability measures within the ADL process risk zone may signal potential for decreased efficiency or safety, and/or increased risk for needing assistance.

Hartman et al. (1999) examined the concept of the risk zone, and of the 6% who were incorrectly classified as either needing or not needing assistance, two thirds of these clients fell above the cutoff but within the risk zone. In their interpretation, this supported the need to also use the risk zone, for while the ADL process ability measures incorrectly classified these individuals as independent (i.e., they had ability measures greater than the cutoff measure), they were in the risk zone and were actually needing some assistance to live within the community.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

As we again consider the idea of objective measurement, it becomes important to realize that not only does the AMPS computer-scoring program utilize a Rasch analysis model to convert raw ordinal data into linear ability measures, but a Rasch measurement model and the philosophical and mathematical principles of modern test theory were also used to guide the development, standardization, and validation of the AMPS. In accordance with the principles of objective measurement and modern test theory, the AMPS has been shown to consist of two unidimensional scales of personal and domestic ADL ability (Fisher, 1993, 1994, 2003a). Additionally, the basic assumptions of a multi-faceted Rasch model of the AMPS have been met, in that (a) all persons are more likely to receive higher scores on easier items/tasks than harder items/tasks, (b) persons of higher ability are more likely to receive higher scores than are persons of lower ability, and (c) strict raters are more likely to assign lower scores across all items than are lenient raters. Such assertions have been verified across different diagnoses (Cooke, Fisher, Mayberry, & Oakley, 2000; Doble, Fisk, Fisher, Ritvo, & Murray, 1994; Doble, Fisk, MacPherson, Fisher, & Rockwood, 1997; Girard, Fisher, Short, & Duran, 1999; Hartman et al., 1999; Kottorp, Bernspång, & Fisher, 2003; Kottorp, Bernspång, Fisher, & Bryze, 1995; McNulty & Fisher, 2001; Oakley, Duran, Fisher, & Merritt, 2003; Pan & Fisher, 1994; Robinson & Fisher, 1999; Robinson & Fisher, 1996; Sellers, Fisher, & Duran, 2001), across various ages, (Dickerson & Fisher, 1993; Hayase et al., 2004), between men and women (Duran & Fisher, 1996; Merritt & Fisher, 2004), and across different cultures (Bernspång & Fisher, 1995b; Dickerson & Fisher, 1995; Goldman & Fisher, 1997; Goto, Fisher, & Mayberry, 1996; Magalhães, Fisher, Bernspång, & Linacre, 1996; Stauffer, Fisher, & Duran, 2001).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Within the examination of the validity of the AMPS and the verification that the basic Rasch assumptions hold true across all persons, tasks, items, and raters, one will find within the literature that there are some minor differences in the item hierarchies between diagnoses (Bernspång & Fisher, 1995a; Cooke et al., 2000; Fisher, 2006a; Rexroth et al., 2005), and between men and women (Duran & Fisher, 1996; Merritt & Fisher 2004). However, it is important to note that these minor disruptions (often statistically significant but not clinically meaningful) do not impact the overall estimation of the ADL motor and/or ADL process ability measures. Thus, the AMPS is considered a valid ADL measure among various diagnoses and between men and women.

To summarize the positive aspects of the AMPS in the realm of objective measurement, I conclude that (a) across persons 3 years of age and older, the AMPS has been shown to consist of two unidimensional scales of ADL ability; (b) the AMPS generates linear, equal interval measures of ADL motor and ADL process ability; and (c) ADL ability measures can be directly compared regardless of the person performing the tasks (sample-free item calibration), the subset of items that were scored, the tasks the person performed (test-free person measurement), or the severity of the rater. What this means is that occupational therapists have the capability of generating objective measures of ADL motor and ADL process ability based on the client performing familiar, relevant, and chosen ADL tasks.

Although commendable, in my opinion a bit more work needs to be done to complete the cycle of objective measurement. That is, in many settings, occupational therapists have not generated what Fisher (2000) terms social objectivity. To complete the cycle, therefore, we need to educate ourselves, our colleagues, and our clients as to

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

what the ADL motor and ADL process ability measures *mean*. For example, if I state that a book is 1 centimeter thick, most readers can instantly generate an idea as to how thick the book is, and most will generate a similar meaning (strong symbol–meaning coordination).

But when I state that a client's ADL process ability was -0.71 logit, occupational therapists and other healthcare professionals are often not able to generate a general idea of the person's ADL ability. Once occupational therapists and their colleagues can generate similar meanings from the linear ADL motor and ADL process ability measures, the AMPS can be considered to be fully objective, for social objectivity will have been achieved. The generation of social objectivity has likely already started within small pockets of rehabilitation personnel, but there is a need to continue the process through published research and clinical education.

If I now come back to the original topic, predicting level of dependence for community living, recall that the AMPS is an assessment of ADL ability, and not one of level of dependence. Moreover, the cutoff measure on the ADL process scale only indicates that the person's ADL task performance is inefficient (time and space organization) and whether or not the person likely needs assistance to live in the community. There is no indication of the amount or type of support that the client likely needs. As a clinician trying to determine and/or justify the need for community support to other healthcare professionals and family members, a scale that yields more information would be useful. For example, clinical markers along the scale that indicate (a) the client is likely able to live independently in the community; (b) the client likely needs minimal support, assistance, and/or supervision to live in the community; or (c) the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

client likely needs moderate support, assistance, and/or supervision to live in the community may be a bit more useful in justifying and documenting a client's future needs. Further delineation of the ADL process scale (and/or ADL motor scale) may also aid in building social objectivity, as clinicians may find it easier to derive a greater sense of what the continuum of ADL ability means within the clinical arena.

Functional Independence Measure (FIM™) (USD$_{MR,}$ 1997)

The Functional Independence Measure (FIM™) is one of the most widely used assessments of ADL used in hospital-based rehabilitation (Lundgren-Nilsson et al., 2005, Månsson & Lexell, 2004). The FIM™ consists of 18 items within the areas of self-care, sphincter control, transfers, locomotion, communication, and social cognition. The above areas have been placed into two domains: FIM™ motor items and FIM™ cognitive items (Table 7). Each item is scored on a scale from 1 to 7, with a 7 indicating complete independence and a 1 indicating the need for total assistance (Table 7). A decision tree format is utilized to aid in scoring each of the 18 items. Upon scoring all 18 items, raw scores are summed to reveal the total FIM™ score or the FIM-18 score; the raw scores for the motor and cognitive items can also be separately summed. The FIM-18 totals can range from 18 to 126 points, with 126 indicating the highest level of independence.

The developers of the FIM™ report that it is an instrument that is used to measure the severity of disability (burden of care), which is defined as "the type and amount of assistance required for a person with a disability to perform basic life activities" (Deutsch et al., 1996, p. 268). Disability is further defined as "any restriction or lack of ability to perform an activity in a manner or within a range considered normal for a person of the same age, culture and education" (USD_{MR}, 1997, p. H-2).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Given that the purpose of the FIM™ is to measure severity of disability, the developers state that the FIM™ is not a measure of impairment, and thus can be used to assess an assortment of individuals with a variety of diagnoses (Heinemann, Linacre, Wright, Hamilton, & Granger, 1993; Stinemann et al., 1996; USD_{MR}, 1997). Furthermore, the FIM™ is a discipline-free instrument, in that any trained clinician (e.g., occupational therapist, physical therapist, nurse, speech therapist, physician) can score the 18 items within the FIM™. Additionally, trained clinicians can share responsibility for scoring the FIM™, focusing on those items related to their specific areas of expertise/knowledge. The FIM™ can be scored via observation of the task performances; through interviewing the client, staff/team members, and/or caregivers; and/or through reviewing the client's medical record (Deustch, Braun, & Granger, 1996). Typically, the FIM™ is administered 72 hours after admission, 72 hours prior to discharge, and 80 to 180 days after discharge via a telephone interview.

The reliability of the FIM™ has been assessed via a meta-analysis of 11 published articles, showing a median inter-rater reliability of 0.95 for the total FIM score, median test-retest reliability of 0.95, and median equivalence reliability of 0.92 (Ottenbacher, Hsu, Granger, & Fiedler, 1996). Researchers have also made judgments regarding the validity of the FIM™, in that the resulting measures can aid in distinguishing higher functioning clients (who need little to no assistance to live in the community) from those who need moderate to minimal assistance to live in the community (Pollak, Rheault, & Stoecker, 1996).

The raw FIM™ totals have also been shown to correlate well with estimated hours of care required (r = -0.76), however the association with estimated hours of

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

supervision was poor (r = -0.39) (Disler et al., 1993). Other studies have documented the association between the FIM™ totals and minutes of assistance provided (Deutsch et al., 1996; Granger et al., 1990, 1993, 1995). In studies with persons with multiple sclerosis, stroke, and brain injury, it was found that a 1-point increase in the FIM™ total score was equivalent to approximately 2.19 to 5.12 minutes of care per day (Granger et al., 1990, 1993, 1995). Additionally, for a sample of individuals within a rehabilitation center due to stroke, a 10-point improvement in the FIM™ total score was associated with a 50% decrease in the amount of care required (Dromerick et al., 2003).

Utilizing a Rasch measurement model, researchers investigated the cross-cultural validity of the FIM™ and found that the number of scoring categories currently used in the FIM™ is not sustainable across a variety of European countries (Lundgren-Nilsson et al., 2005). Furthermore, after collapsing the disordered categories to accommodate for the problem, the items were found to have different levels of difficulty across the various counties. To achieve adequate fit to the Rasch model, the authors noted that a sophisticated adjustment would be necessary to accommodate for disordered thresholds and differential item functioning (Lundgren-Nilsson et al., 2005). As the FIM™ does not utilize a Rasch computer scoring program in the routine use of the assessment, such methods to achieve fit to the Rasch model would likely be impractical for clinical use.

In a similar study, Tennant et al. (2004) also found that when analyzing data from several European countries, that the overall fit to the model of the FIM motor scale was poor and that only 5 of the 13 items had ordered thresholds (e.g., the 7 scoring categories did not show an increasing amount of the construct). As a result, eight of the disordered items were rescored, three of which were converted into dichotomous categories. After

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

rescoring, none of the FIM motor items were found to have differential item functioning by gender or age, but eight items were found to have differential item functioning by country. Following this finding, the authors allowed the items to be unique across the countries, and by doing so, resolved the lack of fit to the model. In summary, the items and the scoring categories of the FIM function differently for different countries, and thus the validity of comparing raw data between countries is questionable.

The evidence of ceiling effects when using the FIM is a bit inconsistent. For example, with a sample of stroke survivors in a rehabilitation setting, researchers found the FIM to be sensitive to changes in disability without demonstrating ceiling effects (Deomerick et al., 2003). In a later study, Lungren-Nilsson et al. (2005) determined, with a sample of hospitalized individuals with stroke, that the range of measurement for the FIM™ motor scores indicated only minor floor and ceiling effects, however floor and ceiling effects were somewhat higher for the FIM™ social-cognitive scores. In contrast to the above findings, when a sample of individuals with traumatic brain injuries, who had received inpatient rehabilitation services 2 to 9 years prior to the study, were tested with the FIM, most of the participants obtained maximal scores on the FIM motor and total FIM scales (Hall, Bushnik, Lakisic-Kazazic, Wright, & Cantagallo, 2001). Likewise, in a clinical study of patients with multiple sclerosis, Månsson and Lexell (2004) found substantial ceiling effects for those individuals who were independent (or nearly independent) in the performance of PADL tasks. Conversely, through utilizing the AMPS, Månsson and Lexell were able to avoid floor and ceiling effects, as the PADL and IADL tasks available within the AMPS offered a wider range of task challenge. Thus, it appears that for higher functioning clients, the FIM has limited ability in

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

measuring functional status and/or functional changes, as significant ceiling effects are likely to occur.

Researchers have utilized admission FIM™ scores to predict discharge setting. Oczkowski and Barreca (1993) found that stroke survivors with total FIM™ scores on admission that were lower than 36 were likely discharged to institutional settings, and those with total FIM™ scores above 96 were discharged back home. Those clients with total FIM™ scores between 36 and 96 on admission changed the most during their rehabilitation stay and the discharge setting was not predictable. As there is a large gap between those with a total FIM™ score of 36 and 96, it seems prudent to question the sensitivity of the FIM™ in predicting discharge setting.

Lastly, Granger et al. (1993) found that with a sample of patients with stroke that 30% of those with discharge FIM™ scores less than 40 returned to the community, 67% of those with discharge FIM™ scores between 60 and 79 returned to the community, and 86% of those with FIM™ scores of 80 or higher returned to the community. In a more recent study with 37 patients who had sustained burn injuries, total FIM scores at the time of discharge that were greater than 110 points were associated with discharge back home, whereas total FIM scores that were lower than 110 points were associated with the need for continued rehabilitation (Choo, Umraw, Gomez, Cartotto, & Fish, 2006). Although researchers have found trends with regard to predicting discharge setting, there appears to be inconsistent findings across the different client populations. A discharge total FIM score of 80 or higher was indicative of a discharge home within a sample of individuals with stroke, whereas a discharge total FIM of 110 was indicative of a discharge home for those who had sustained burn injuries. Are the inconsistent findings due to differences in

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

client characteristics or are they due to properties of the FIM? Further research will need to occur to answer such questions.

The first critique of the FIM™ pertains to the claim within the user manual that it does not measure impairment (i.e., body function). Through looking at the specific scoring examples, this statement may be misleading. For example, as one looks at the two items regarding bladder and bowel management, it becomes quite clear that sphincter control is at the root of these items. Sphincter control is a medical issue and is not related to the construct of disability. When scoring these items, however, the client must maintain continence (e.g., sphincter control) at all times, without the use of equipment or agents, to be scored as being independent in this area. If clients were solely scored on their ability to manage bladder/bowel difficulties (e.g., through the use of undergarments, catheters), then one could consider this to be within the realm of disability and/or occupational performance. As a result, the relative difficulty of these items will depend solely on one's medical condition (i.e., the presence or absence of sphincter control), and thus some individuals will have a disadvantage (or advantage) on these items based only their medical diagnoses. Additionally, bladder and bowel management often fail to demonstrate adequate fit to the model, indicating that these items may not contribute to a unidimensional scale (FIM™ motor).

Additionally, when one examines the item of comprehension, those with certain medical conditions (e.g., diminished hearing and/or vision) are at a disadvantage. In order to obtain a score of independence on this item, one cannot utilize a hearing aid and/or glasses to comprehend visual or auditory information. Given this, many individuals (e.g., older adults, individuals with developmental disabilities) will score

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

lower on this item secondary to their diminished body functions. Lastly, memory is not clearly an item reflecting occupational performance, for according to the scoring criteria, "memory in this context includes the ability to store and retrieve information, particularly verbal and visual" (USD_{MR}, 1997, p. 47). Although the authors preface that this item is scored with regard to the client's daily task performances (especially remembering people, daily routines, and requests), it is unclear how this item can be differentiated from the body function of memory. Thus, it appears that several items within the FIM™ may measure at the impairment level, and not solely at the functional or occupational level.

Given this, it is important to note that while the scoring criteria of the FIM™ is primarily based on the need for assistance, it appears as though some of the items may be linked to certain medical diagnoses. If this is the case, bias may be present when administering this assessment to different client populations. An additional literature search revealed some information with regard to this issue. For example, one study examined the reliability and validity of the FIM™ for persons 80 years of age and above (Pollak et al., 1996). The researchers subjected the data to Rasch analysis, revealing the presence of two unidimensional scales (FIM™ motor and FIM™ cognitive). When they measured how well the items fit the constructed model, it was found that bladder management, bowel management, and grooming misfit the model on the FIM™ motor scale and memory misfit the model on the FIM™ cognitive scale. When the item calibration values (a measure of relative item difficulty) of this older sample were compared to younger sample, they found a high correlation ($r = 0.9$) between the motor item calibration values and a low correlation between the cognitive item calibration values ($r = -0.3$).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Due to the operational definition of bladder/bowel management as being continence, many individuals in the older adult study were scored down on these items (Pollak et al., 1996), and thus they experienced this item as being relatively harder than did the younger sample. More specially, the older sample found these items harder to endorse secondary to poor sphincter control. Similarly, items such as problem solving and comprehension were relatively harder for the sample of older adults. Problem solving calibrated as being harder within the sample of older adults because lower scores were obtained on this item due to the fact that many of these individuals delegated financial responsibilities to family members/custodians. Comprehension also calibrated as being relatively harder for the sample of older adults, for many of these individuals utilized visual and/or hearing aids.

Another study investigated the structure and stability of the FIM™, and also found that the FIM™ items define two statistically and clinically different indicators (FIM™ motor and FIM™ cognitive) (Linacre, Heinemann, Wright, Granger, & Hamilton, 1994). It was found that the two distinct measures of disability offer more useful information than the combined measure. However, in their investigation, the bowel and bladder items again failed to fit the model, and suggestions were made to distinguish continence and managing episodes of incontinence into two separate items. With adaptations to the FIM™ motor items, the authors note that the assessment would produce an even more precise and useful measurement system. In looking at the test manual (USD_{MR}, 1997), it appears as though such suggestions were not implemented, as the bowel and bladder items still revolve around continence (vs. management of episodes of incontinence).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

In an earlier study, Heinemann and colleagues (1993) determined that the FIM™ could be used in a valid manner across an assortment of diagnostic groups. However, they did note that the item difficulty calibration values did vary across diagnostic groups. Within their findings, they concluded that the motor scale could be used for all impairment groups, except those with back pain and burns. Likewise, one cognitive scale could be used for all individuals except those with stroke, brain dysfunction, and congenital impairments. A separate cognitive scale was developed for the latter three diagnoses. Such findings are somewhat alarming, and the need for further investigation of the effect of these disruptions on the overall score is definitely needed.

Although the FIM™ was not developed utilizing Rasch measurement methods, researchers have subjected FIM™ data to Rasch analysis during their continued investigation of the tool. It is important to note that within the family of Rasch measurement models, bias occurs when there is a significant and persistent interaction between some (but not all) persons and some (but not all) items/tasks (Wright & Stone, 1988). Thus, bias may be present if one population has an unfair advantage (or disadvantage) when performing some, but not all items/tasks within the assessment. Thus, a primary assertion has been violated, in that the FIM™ has not been found to exhibit sample-free item calibration. What all of this boils down to is the fact that the FIM™ may function differently for different client populations/cultures and diagnoses. Within this, it is questionable if the items within the FIM™ truly define two unidimensional scales (FIM™ motor and FIM™ cognitive) that measure disability and/or burden of care. Considering that the FIM™ may not truly define two unidimensional scales of disability across assorted diagnostic groups and world regions, and the potential

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

problems of floor and ceiling effects, the FIM™ is a less than ideal assessment of disability.

Lastly, the utilization of raw, ordinal data within objective measurement is considered erroneous (Bond & Fox, 2001; Smith & Smith, 2004; Wright & Masters, 1982; Wright & Stone, 1979). Furthermore, given that the FIM™ functions differently among different client populations (e.g., different ages, diagnoses, world regions), comparing FIM™ raw scores is especially questionable and limited in scope. In comparison, although researchers have found minor disruptions in the AMPS item hierarchies between client populations, the minor disruptions did not impact the estimation of the ADL motor or ADL process ability measures (Bernspång & Fisher, 1995a; Cooke et al., 2000; Duran & Fisher, 1996; Fisher, 2006a; Merritt & Fisher 2004; Rexroth, Fisher, Merrrit & Gliner, 2005). Thus, researchers and clinicians can validly compare AMPS ADL motor and ADL process ability measures across and within different client populations.

The Functional Autonomy Measurement System (SMAF)

The Functional Autonomy Measurement System (SMAF) (Hébert et al., 1988; Hébert, Guilbault, Desrosiers, & Dubuc, 2001) was designed to measure disability and handicap, as defined by the World Health Organization classification of disablement (ICIDH) (WHO, 1980). Please note that since this assessment was developed before the most recent update to the *International Classification of Functioning, Disability and Health* (ICF) (WHO, 2001), the terminology is outdated. Although the terminology of the assessment is outdated, the items have been framed to capture disabilities that are related to the interactions between individuals and their environments (Dubuc et al.,

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

2005), which is consistent with conceptual foundations ICF (WHO, 2001). However, as a means of being consistent with the authors of the SMAF, I will use the older terminology of the ICIDH that was used within their research articles. Thus, disability was defined as diminished functioning in daily activities as a result of impairment. Likewise, handicap implied a social disadvantage, resulting from the demands which were imposed on the individual and the material and social resources available to overcome one's disability (Hébert et al., 1988).

The SMAF consists of 29 items that cover the following five subscales: ADL, mobility, communication, mental functions, and instrumental activities of daily living (IADL) (Hébert et al., 1988, 2001) (Table 8). Additionally, researchers have begun the preliminary work to include an additional subscale of social functioning (Pinsonnault et al., 2003). The SMAF is typically completed by a healthcare professional (e.g., nurse, case manager, occupational therapist) who has either interviewed the client and/or caregiver(s), and/or observed the client's performance. When scoring each item, the examiner scores the individual on the amount of assistance the person needs to carry out the activity (disability score) and the resources that are available to the individual during the performance of the activity (handicap score). Lastly, the examiner notes who is available to provide assistance and the stability of this assistance (e.g., if the caregiver is able to continue to provide such assistance). Disability is scored on a 4-point scale: zero (independent), -1 (needs supervision), -2 (needs help), and -3 (dependent). In the revised scale, the ADL, mobility, and IADL items are scored on a 5-point scale, with the addition of -0.50 indicating that the individual performs the activity independently, but with some difficulty (Desoriers, Bravo, Hébert, & Dubuc 1995).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

One of the greatest strengths of the SMAF is that it can be used to assess both disability and handicap. As a result, an examiner may be able to obtain a better picture of one's functional needs. For example, when assessing a client's ability to negotiate stairs, a score of -3 could be assigned, indicating that the client cannot negotiate stairs. However, when scoring this item within the context of the client's actual residence, one may determine that the client does not have stairs, and thus, there is no immediate need for him/her to be able to perform this task. The combination of the two scores may aid examiners in creating a realistic, client-centered picture of a client's actual need for assistance.

As a means of comparison, when administering the AMPS, if the client cannot be assessed within his/her own home, every attempt is made to make the environment as relevant and similar to that which would be encountered within the home. However, when administering the AMPS, the examiner does not directly address the available assistance or the stability of the available assistance. Within the FIM, such issues are not taken into consideration at all, for if a client cannot negotiate stairs, he/she is simply scored as needing total assistance in that area.

Hébert and colleagues (1988) determined that the SMAF has acceptable inter-rater reliability, within and across different professions and settings (e.g., nursing, social work; community, institution), with a mean agreement of 75% for all of the items on the scale and a weighted kappa ranging from 0.52 to 0.76 ($p < 0.01$). Additionally, acceptable concurrent validity was demonstrated upon finding a strong correlation between the SMAF and the amount of required nursing-care time (Hébert et al., 1988). An interesting point with regards to this study is that items with the highest correlation to

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

time of nursing care needed were the ADL items ($r = 0.89, p < .001$). Furthermore, the items with the lowest correlations were the communication and mental functioning items ($r = 0.58$ and $0.63, p < .001$ respectively).

Desrosiers, Bravo, Hébert, and Dubuc (1995) conducted a follow-up study to investigate the reliability of the revised SMAF (the addition of a 5-point rating scale for selected items). Within this study, they found the test-retest and inter-rater reliability (intra-class correlation [ICC] = 0.95 and 0.96 respectively) to be acceptable for clinical and epidemiological purposes. With respect to the addition of the social subscale, after minor revisions (from version 1.0 to 1.1), the interrater reliability of the social subscale showed moderate to almost perfect agreement, with the weighted kappas ranging from 0.38 (social relationship item) to 1.0 (social network item) and an ICC for the social functioning subscale of 0.83 (Pinsonnault et al., 2003).

To investigate the validity of the SMAF, researchers have compared the SMAF with other ADL assessments. Researchers have documented that the SMAF total scores are highly correlated to total FIM scores (r ranging from 0.93 to 0.95) and the Barthel Index ($r = 0.92$) (Desrosiers et al., 2003; Langlais as cited in Hébert et al., 2001). The high correlations found with similar ADL assessment scores documents the concomitant construct validity of the SMAF. Additionally, SMAF scores have been indirectly compared to AMPS ADL ability measures. More specifically, researchers sought to determine how much variance in functional autonomy (SMAF scores) and functional performance (AMPS ADL ability) is explained by body functions (i.e., motor factor [upper extremity and balance], cognitive factor, and perceptual factor) (Mercier et al., 2001). Results indicate that the AMPS ADL ability measures explain only 64% of the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

variance in the body function variables, whereas the SMAF explains 93% of the variance in the body function variables. Results can be viewed in two ways; the authors note that little is gained by utilizing assessments of body functions in conjunction with the SMAF. However, one may also conclude that the SMAF scores are more linked to body functions than true performance. Within this line of reasoning, one could also conclude that the AMPS stands out as the evaluation that assesses at the level of participation and occupational performance (e.g., the actual doing of relevant daily life tasks vs. body functions). Thus, within a battery of assessments that are utilized, one may consider using both the SMAF (assessment of functional autonomy/handicap) and the AMPS (assessment of the quality of ADL performance), as the two assessments likely document somewhat different aspects of function.

In looking at responsiveness to change, research has shown that a 5-point change in the client's total SMAF score corresponds to a metrically and clinically significant change (Hébert, Speigalhalter, & Bayne, 1997). In comparing the responsiveness of the SMAF with other common assessments of ADL ability, Langlais (as cited in Hébert et al., 2001) found that the SMAF, FIM, and Barthel Index are statistically similar with regard to responsiveness to change. More recently, researchers compared the responsiveness of the SMAF and the FIM and found that corresponding categories within the FIM and the SMAF are similarly responsive/sensitive, although the total SMAF scores were more responsive than the FIM total scores (Desrosiers et al., 2003). The difference in the responsiveness of the total scores is likely due to the inclusion of an IADL category on the SMAF, with no corresponding category found in the FIM.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Lastly, SMAF total scores have been utilized as a means of classifying older adults into disability profiles (ISO-SMAF profiles) (Dubuc et al., 2005). The characteristics of the individuals in the 14 different profiles range from (a) needing supervision for housekeeping and difficulty with meal preparation, transportation, and budgeting (Iso-SMAF profile 1, mean score of 9.33); (b) having moderate alterations in cognitive functions, minor behavioral problems, difficulty with ADL, and the need for assistance with washing and grooming (Iso-SMAF profile 7, mean SMAF score of 39.19); and (c) being bedridden and dependent in ADL (Iso-SMAF profile 14, mean SMAF score of 73.77). The results supported the use of the Iso-SMAF profiles in determining not only the level of care needed by elderly clients, but also within the determination of the monetary and staff requirements of residential care facilities (Dubuc et al., 2005; Tousignant, Hébert, Dubuc, Simoneau, & Deileman, 2003).

The creation of the 14 Iso-SMAF profiles is a step toward the utilization of an ADL assessment in the determination of level of dependency. One limitation of the profile system is that has been developed and utilized only with older adults. Thus, there is no indication as to whether or not the profiles would remain valid with a more diverse population. Secondly, a few quick calculations revealed that the mean difference between adjacent Iso-SMAF profile mean scores is -4.96 points (ranging from -0.24 to -8.79), and the mean standard deviation of the Iso-SMAF profile scores is 4.56 (ranging from 3.58 to 5.93). Thus, within a reasonable range of potential error (i.e., within ±2 standard deviations), there is a great deal of overlap in SMAF scores for the 14 profiles. For example, to be considered within Iso-SMAF profile 1, an individual could reasonably have a SMAF score between 2.17 and 16.49 and between 4.84 and 21.61 to be considered

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

within profile 2 (Tousignant et al., 2003). Thus, it appears as though the categorization scheme may have some limitations with regard to accurately determining the level of disability. Lastly, one must also consider the fact that the SMAF only utilizes a rating scale, and thus ordinal rankings are generated, not equal interval measures. The generation of ordinal ratings within objective measurement is considered erroneous (Wright & Masters, 1982; Wright & Stone, 1979), as it is difficult (if not impossible) to accurately compare summed ordinal ratings.

Index of Independence in Activities of Daily Living (Katz Index of ADL)

The Index of Independence in Activities of Daily Living (Index of ADL), often referred to in the literature as the Katz Index of ADL, is a simple, quick summary of one's ability to perform PADL tasks (Katz et al., 1963, 1970. More specifically, the Katz Index of ADL can be used to assist the examiner in summarizing a client's functional abilities within the following six PADL areas: (a) bathing, (b) dressing, (c) going to the toilet, (d) transferring, (e) continence, and (f) feeding (Katz et al., 1963, 1970). Although the scoring criteria are different for each item, in general, clients are scored as being independent, needing minimal assistance, or needing moderate assistance in their performance of the activity. Upon completing the scoring of the six items, the examiner determines a summary score, or Katz Index of ADL level. The Katz Index level can be A, B, C, D, E, F, or G, where A is the most independent grade and G is the most dependent overall grade (Table 9). This tool has been shown to have appropriate reliability and validity within a short-term care setting (Brorsson & Åsberg, 1984).

The Katz Index of ADL levels have been concurrently correlated with amount of personal assistance required, in that a grade of A correlated with independent functioning

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

within the community and scores of E, F, or G were correlated with needing nonfamily attendant care and a higher rate of nursing home placement (Katz et al., 1970). Additionally, the Katz Index of ADL levels have also been shown to be useful in predicting future assistance needs (Katz et al., 1963). In this study, 79% of those that received a Katz Index of ADL grade of D, E, F, or G at discharge were receiving nonfamily attendant care 1 year after the initial onset of their stroke.

Researchers have also demonstrated that this tool has value in predicting length of hospital stay, type of discharge setting, and survival or death during the first month after experiencing an acute stroke (Åsberg & Nydevik, 1991; Brorsson & Åsberg, 1984). When the data were dichotomized (Katz Index levels ABC vs. DEFG), the two categories were able to predict discharge home versus in hospital or dead 1 month after stroke (sensitivity = 83, specificity = 97) (Åsberg & Nydevik, 1991). Although such predictive values are high, there is no indication regarding the amount of assistance the individuals needed in order to return home. Those with a Katz Index of ADL score of A had shorter hospital stays, were likely discharged home (88%), and were more apt to be living at home 1 year after discharge (72%). In contrast, those with Katz Index of ADL scores of E, F, or G had longer hospital stays, were more likely to be discharged to institutional care (45%), or were more likely to die (26%), and most were living in institutions (26%) or were dead (54%) 1 year after discharge (Brorsson & Åsberg, 1984).

In terms of sensitivity, Brorsson and Åsberg (1984) commented that within their study, many participants were classified as having either a Katz Index level of A or within the grades E, F, or G, indicating that the Katz Index of ADL likely has limited

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

discriminatory power. As a result, they suggest the use of 3 grades for screening purposes (e.g., A, BCD, and EFG).

Upon investigating the tool, it appears that the three scoring levels used to score each item would not detect small changes in function. For example, when scoring the item of transferring into/out of the tub/shower and washing one's entire body, a client can be scored as (a) independent, (b) needing assistance to wash one part of the body, or (c) needing assistance to wash more than one part of the body. Based on personal clinical experience, there is a great deal of change that can occur in terms of the client's performance that may not be detected using these scoring criteria. For example, a client could go from needing physical assistance with washing more than one part of his/her body to needing only supervision or verbal cues during the bathing sequence. Within the scoring of the Katz Index of ADL, this client would receive the lowest score for bathing for both time one and time two. Thus, the scores attained by this client would not change, even though the client's actual level of care has decreased to a level that the family could easily provide at home.

Lastly, as the items of Katz Index of ADL only address PADL tasks, ceiling effects may result with higher functioning clients. Many higher functioning individuals may be able to independently perform the PADL tasks defined in the Katz Index of ADL, yet not be able to independently perform harder tasks, IADL tasks (e.g., cooking, home maintenance). When using the FIM™ to assess higher functioning clients, ceiling effects were detected in higher functioning clients, as only PADL items were scored (Månsson and Lexell, 2004). Thus, one would anticipate similar findings upon further investigation of the Katz Index of ADL.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Klein-Bell ADL Scale

The last assessment I will review is the Klein-Bell ADL Scale (Klein & Bell, 1982). This assessment tool was designed to assess one's ability to perform PADL tasks in the areas of dressing, bathing/hygiene, elimination, functional mobility, eating, and emergency communication. The tool consists of 170 items, each of which define small, finite steps within each ADL task (Klein & Bell, 1970; Law & Usher, 1988) (Table 3). The client is scored as achieving the item (item performed independently), failing the item (needs assistance to perform the item), or N/A (only if that item is permanently not applicable to this client). The client is not scored down for using adaptive equipment as long as the client is able to perform the item independently. No points are given if the client cannot perform the item independently.

It is important to note that since clients are only scored as either independent or needing assistance, small changes in performance may not be detected, even though multiple finite steps are scored. For example, during the initial assessment a client may be independent, yet experience increased effort and/or decreased efficiency during the performance of the task step. Upon reevaluation, the client may experience less effort and greater efficiency, however the score on the Klein-Bell would not reflect these important changes.

After the examiner has scored all 170 items, each item is multiplied by a weighted value, and these values are then summed within each global area (i.e., dressing, eliminating, mobility, grooming/hygiene, eating, and emergency telephone use) and placed on a graph for an easy visual depiction of the client's abilities. The weighted values were developed according to rehabilitation professionals' judgments regarding the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

relative difficulty of each item. Based on such ratings of relative difficulty, each item is weighted as 1, 2, or 3. A weighted value of 2 indicates that the item fell within 1 standard deviation of the mean ratings given for each item by the rehabilitation professionals. A weighted value of 1 indicates that the item fell more than 1 standard deviation below the mean ratings, and a score of 3 indicates that the item was more than 1 standard deviation above the mean ratings. The sums of each ADL task are then added to yield an overall ADL independence score.

Inter-observer reliability of the Klein-Bell ADL scale is 92% agreement (Klein & Bell, 1982). Furthermore, scores from the Klein-Bell ADL scale have been correlated with the hours per week of assistance one receives with ADL tasks ($r = -0.86$, $p < .001$), in that the lower the score on the Klein-Bell Scale, the greater the amount of assistance needed (Klein & Bell, 1982). Lastly, the scale has been shown to be sensitive to change over time (Shillam, Beeman, & Loshin, 1983).

The Klein-Bell offers a potentially sensitive assessment that is based on the observation of discrete steps within basic ADL tasks. The scale offers healthcare providers valuable information regarding which steps within the task sequence are more difficult for the client, and thus, well targeted intervention strategies could likely be developed. Additionally, the developers have aimed to avoid the mistake of simply summing ordinal ratings, for they have incorporated the use of weighted item values. Although commendable, the evidence within the literature fails to verify the accuracy of the weighted item values. For example, were the rehabilitation professional's judgments regarding the relative difficulty of each item accurate? Are the harder items harder for all individuals, regardless of age or diagnosis? Likewise, are the easier items easier for all

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

individuals, regardless of personal characteristics? Lastly, weighting the items based on a 3 point scale appears to be a bit limiting, for items with the same weighted values may in fact have different levels of difficulty. Thus, although raw scores are not utilized, the methods used to determine the weighted values are less than ideal and have little face value, especially within the realm of modern measurement.

Summary of ADL Assessments

Of the ADL assessments that were investigated, an assortment of potential problems were illuminated. First, the majority of the assessments are based solely on the use of raw, ordinal data. Such practice "violates" the basic premises of modern test theory, problems of which I discussed earlier. Second, many of the assessments were not appraised to offer high levels of sensitivity, as ceiling and floor affects likely occur, and small changes in performance may not be detected (Table 11). In their discussion of the methodological problems of many ADL assessments, Law and Usher highlighted similar findings (1988).

The larger goal within this project is to investigate the validity and reliability of using the AMPS to predict level of dependence. The AMPS offers a wide assortment of benefits including that (a) modern test theory has been utilized in development and standardization process; (b) linear ADL ability measures are generated, and thus direct comparisons between individuals can be made (regardless of which tasks were performed); (c) it has been shown to be a sensitive outcome measure, without the presence of ceiling and floor effects; (d) an established cutoff measure has been developed that can be used to identify clients who likely need assistance to live in the community; (e) it can be utilized with nearly any person 3 years of age or older with any

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

diagnosis; (f) it is a performance-based assessment, not a questionnaire about one's performance; and (g) it is a client-centered, occupation-based analysis of the quality of one's ADL task performance.

Assessments of Independent Living

The next step is to find a gold standard assessment that can be used as an external criterion of one's ability to live independently within the community. Within this investigation, Dever and Knapczyk's (1988) taxonomy of independent living skills will serve as a guide in determining if *all* areas of independent living have been addressed. More specifically, the desired gold standard should meet the following criteria: (a) items within the assessment should cover all or nearly all of the domains of independent living, (b) use of the assessment has been shown to be reliable and valid for use with individuals of varying ages and diagnoses, and (c) the assessment must be shown to be a valid and reliable indicator of the level or amount of assistance that individuals need to live in the community. The following section is a review and critique of some of the available assessments of independent living. Not all of the available assessments of independent living have been reviewed, however a representative sample has been included.

Although the following assessments are used to assess community living skills, interestingly, many appear to primarily assess ability within the domains of PADL and IADL. Thus, there is significant amount of overlap between the ADL scales previously discussed and the independent living skills assessments discussed in this section. This once again lends to the notion that ADL ability may explain a significant amount of the variation in the ability to live independently in the community.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Community Living Skills Assessment Inventory

The Community Living Skills Assessment Inventory (Switzky, Rotatori, & Cohen, 1978) was designed to delineate "an individual's functional living skills in areas considered essential for successful placement in a range of semi-independent community living facilities" (p. 1336). More specifically, this assessment was primarily designed to be used with individuals with developmental disabilities as a means of better understanding each person's independent living skills. Through gaining this knowledge, caregivers can choose the most appropriate living environment, one that provides the individual with the just right challenge and necessary supports. This assessment has also been used to establish training/therapy goals and to analyze the efficacy of skills training programs.

The Community Living Skills Assessment Inventory can be used to assess the following eight areas of function (a) dressing and undressing, (b) personal hygiene and grooming, (c) eating, (d) housekeeping, (e) preparing food, (f) self medication, and (h) functional/adaptive equipment (Table 11). Each performance area is broken down into smaller tasks and task steps. For example, within the area of personal hygiene and grooming, the task of brushing one's teeth has been broken down into the following four steps: (a) open toothpaste, (b) apply toothpastes, (c) rinse mouth, and (d) wipe mouth.

Almost all of the items within the assessment are observable goal-directed sequences. However, within the task of taking one's medication, there are three items that an assessor could not directly observe (a) *knows* when to take his/her medication, (b) *knows* what pills to take, and (c) *knows* the number of pills to take. These three items involve the concept of knowledge, a cognitive function that cannot be directly observed.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

An examiner can only observe on that given day whether or not the person takes the correct medicine at the right time, he/she cannot *directly observe* the person's knowledge or understanding of the correct prescription. Perhaps the item simply needs to be reworded so that it reads that the individual (a) takes the medication at the correct time, (b) takes the correct medication, and (c) takes the correct dose/amount.

The Community Living Skills Assessment Inventory is scored within the client's natural environment by an individual who is most familiar with the client (e.g., teacher, paraprofessional, occupational therapist, psychologist, nurse, volunteer, physical therapist). The assessor scores the client after directly observing the client perform the given task. Performance is timed and scored on a 7-point scale, however it is not clear if each task is scored and timed (e.g., brushing teeth), if each item (e.g., task step) is scored and timed, or if both are scored and timed (Switzky et al., 1978) (Table 13). Since the user manual was not readily available, I was unable to verify what is scored and timed.

Information could not be found via a literature search related to the validity of the Community Living Skill Assessment Inventory for predicting level of dependence. Based on the acquired information, it appears as though clinicians use this assessment at a more "local" level. Thus, clinicians may view the client's scores and overall profile and compare this to the available housing options in their area, seeking to create a just right challenge within the client's living environment.

As one reviews this assessment, several things come to mind. First, the researchers have claimed that "the inventory measures the independent quantitative functional living skills . . . in eight major skill areas in actual living situations" (Switzky et al., 1978, p. 1341). If, within this statement, the researchers have proposed that one

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

generates quantitative measures of skill/ability, then the researchers have misinformed their readers, for a truly quantitative/objective measure is one that is equal interval (Bond & Fox, 2001; Smith & Smith, 2004; Wright & Masters, 1982; Wright & Stone, 1979). That is, the utilization of a 7-point rating scale generates ordinal data, not equal interval/quantitative data. Although time is an equal interval measure of the amount of time spent performing tasks, it is not an equal interval measure of actual community living skill. For example, a client could perform a task very quickly, but he/she may experience inefficiency, effort, the need for assistance, and/or decreased safety. Thus, within this measurement there is a lack of information about the quality of the client's skills within the performance of every day tasks.

Overall, this assessment may meet the needs of some professionals who are seeking to determine an individual's overall ability to perform the selected PADL and IADL. The acquisition of such knowledge could potentially aid professionals and caretakers in the creation of a training program for the development/enhancement of PADL and IADL skills and/or the evaluation of such a training program. However, this assessment tool would not be an ideal outcome measure, as the small changes in the client's performance may not be captured. For example, during an initial evaluation, a client may receive a rating of 5 (independent but imperfect performance) when cooking a meal. During this initial evaluation, the examiner may observe marked effort and inefficiency during the task performance. Upon reevaluation, the client and therapist note improvement (e.g., client experienced minimal effort and minimal inefficiency), but errors of performance were still observed and thus the client once again obtained a score of 5 on the cooking task. In this situation, although the client has shown a clinically

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

meaningful improvement, the assessment was not sensitive enough to reveal the changes in performance.

Given that one of the uses of the results of this assessment is the determination of the most appropriate community setting (e.g., group home) for individuals with developmental disabilities, it seems that social interaction and community mobility would also be important aspects to assess. Clearly, individuals must effectively interact with others (e.g., housemates, staff) if they are to successfully live within such settings. Additionally, such individuals must possess the ability to travel within one's home and within and between community settings. Thus, it seems logical that such items would be important to assess, especially given the intended use of the results of this assessment.

In addition to the above noted tasks/skills, there are many other areas of independent living that are not assessed within the Assessment of Community Living Skills Inventory. For example, tasks such as purchasing supplies, managing one's finances, caring for the outside of one's home, and leisure have not been included within this assessment. Thus, although reported as an independent living skills assessment, it is appears as thought this is an assessment of one's ability to independently perform a subset of PADL and IADL tasks.

St. Louis Inventory of Community Living Skills

The St. Louis Inventory of Community Living Skills (SLICLS) (Evenson & Boyd, 1993; Fitz & Evenson, 1995) was designed to assist healthcare providers in determining the most appropriate community living setting for individuals with mental illness. The SLICLS, consisting of 15 items, can be completed in 2 to 3 minutes. Each item is scored on a 7-point scale by a rater who is familiar with the client (e.g., caregiver,

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

healthcare provider, social worker) (Table 14). The rater assigns a score for each of the 15 items based on the client's current performance, using the following criteria: a score of 1 indicates that the client possesses few or no skills, a score of 2 indicates that the client needs substantial improvement or help, a score of 3 indicates the need for moderate improvement or help, a score of 4 indicates the skill level was not clear or not observed, a score of 5 indicates that the client has moderate skills, a score of 6 indicates that the client has substantial skills, and a score of 7 indicates that the client is self-sufficient (Evenson & Boyd, 1993). Note that an average rating of 4 is assigned if and when the rater does not have the opportunity to observe the skill or if the skill level is not clear. The authors failed to explain this rationale, thus it is unclear as to why this rating is assigned. The researchers have only noted that an average rating is the most "accurate" score for such instances.

Authors have documented that the SLICLS has acceptable inter-rater reliability, with an intra-class correlation coefficient of .74, and acceptable internal consistency, with an alpha of 0.97 (Fitz & Evenson, 1995). Although the given correlations of internal consistency are acceptable within many research circles, such statistics tend yield a reliability estimate which "misleads the test-user into believing a test has better measurement characteristics than it actually has" (Linacre, 1997, p. 581).

The validity of the SLICLS scores was supported by the finding that the mean total scores differed significantly between groups of individuals from three different residential settings ($F = 32.9, p < .0001$). Additionally, all three mean scores differed significantly from each other. As was expected, individuals living in apartments had the highest mean scores (80.38, $SD = 12.64$), individuals living in nursing homes had the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

lowest mean scores (41.4, *SD* = 20.12), and individuals living in residential care facilities fell in between these two means (68.78, *SD* = 12.75) (Fitz & Evenson, 1995). Furthermore, the SLICLS significantly correlated with the social worker's estimate of the best level of placement ($r = .65$, $p < .001$).

In a study examining the validity and reliability of a Chinese version of the SLICLS (SLICLS-C), the authors document that the results can be utilized to predict level of community care in a group of Chinese adults with schizophrenia (Au, Tam, Tam, & Ungvari, 2005). Through examining the sensitivity and specificity of various cutoff scores, the researchers chose the most optimal cut points to differentiate those who were independent in the community, semi-independent in the community, and in dependent on others to live in the community. A SLICLS-C score of 85 (sensitivity = 0.90, specificity = 0.81) was chosen as the cut point for independent versus semi-independent and a score of 66 (sensitivity = 0.91, specificity = 0.55) was chosen as the cut point for semi-independent versus dependent (Au et al.). From this study, one can see that the assessment may be more accurate when predicting level of community care for those who are more able, as the those within the dependent sample were correctly categorized only 55% of the time (vs. 90% for those who are independent and between 81 and 91% for those who are semi-independent).

Although the researchers have concluded that the SLICLS has adequate reliability and validity, the astute reader may develop alternative conclusions. First and foremost, although statistical differences were found between the three different types of residential settings, the large standard deviations associated with the mean scores of the three different levels of care were quite high, indicating that there was a great deal of

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

variability among the individuals within each group. In fact, the range of expected scores, ± 1 standard deviation, for those living in the nursing home was between 21.28 and 61.52 (or between 1.16 and 81.64 based on ± 2 standard deviations). Such a range of ability seems very large, and if one were to assume that all of the individuals living within the nursing home truly needed a high level of assistance (within an acceptable range of variability), doubt arises as to the accuracy/validity of the results of this assessment. Lastly, researchers (Au et al., 2005; Fitz & Evenson, 1995) chose to compare and sum raw ordinal data. As discussed previously, such practices within objective measurement are considered erroneous (Bond & Fox, 2001; Smith & Smith, 2004; Wright & Masters, 1982; Wright & Stone, 1979).

Independent Living Skills Survey

The Independent Living Skills Survey (ILSS) (Wallace, Boone, Donahue, & Foy, 1985) was designed to provide "detailed assessment of the community living skills of chronically mentally ill patients residing with significant others or in residential care facilities" (Cyr, Toupin, Lesage, & Valiquette, 1994, p. 92). Within this assessment, 9 to 10 different areas of independent living are assessed via the use of one of two versions of the assessment. Within the significant-other version, a family member or caretaker fills out the survey, and within the patient version, the client fills out the survey (i.e., self-report version). In the self-report version, 10 areas of living skills are assessed, including (a) personal hygiene, (b) personal appearance and care of clothing, (c) care of personal possessions, (d) food preparation and storage, (e) health maintenance, (f) money management, (g) transportation, (h) leisure and recreation, (i) job-seeking skills, and (j) job maintenance skills. There are 65 items that make up the assessment, each of which

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

falls under one of the above noted areas of independent living. Although a complete listing of the items was not included within the research article, some item examples include: (a) Do you buy your own clothing?; (b) Do you bathe or shower using soap at least twice per week?; (c) Do you eat at least one serving of a vegetable daily?; (d) Do you administer your own medications?; (e) Do you pay your own bills such as rent, utilities, and phone?; and (e) Do you have a hobby on which you work regularly? Each item is scored on a 3-point scale, with a score of 1 indicating that the client performs the item, a score of 2 indicating that the client does not perform the item, and a score of 3 indicating that the item is not applicable (Cry, Toupin, Lessage, & Valiquette, 1994).

Within the investigation of the reliability and validity of the self-report version of the ILSS, the researchers concluded that the instrument has adequate psychometric qualities (Cry et al., 1994). For example, after the exclusion of six items, which had either too high or too low item-total correlation or variance, the internal consistency of the items (Cronbach's alpha) ranged from 0.47 to 0.93, with a mean of 0.69. The lowest coefficient was for the leisure and recreation scale, and the highest coefficient was for the job-seeking scale.

Cry and colleagues (1994) also concluded that the ILSS has adequate construct validity, as evidenced by the fact that groups of individuals from various settings (short term ward, independent apartment, hostel, foster family) had significantly different total scores within all but 1 of the 10 scales of ILSS. Although there were four different groups from four different residential settings, the authors failed to specify which groups differed; they merely cited that the groups differed significantly. Thus, based on their report, the reader does not know if one group differed from the other three groups, if all

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

the groups differed, or if only the most extreme groups differed (e.g., short term ward vs. independent apartment). The researchers also commented that the scales within the ILSS demonstrate construct validity, for similar to other studies, women were significantly more able than men on 4 of the 10 scales (personal hygiene, personal appearance, care of personal possessions, and food preparation and storage) and on the total score.

As one critiques the work of Cry et al. (1994) in their investigation of the ILSS, many things come to mind. First of all, the authors, like many other authors, chose to sum and compare raw ordinal data, and as noted above, such practices are questionable. Second, it was curious that the researchers chose to compare the differences that existed between men and women as a means of establishing construct validity. According to Cry et al. (1994), others have found that women are more capable of living autonomously within the community than are men, and thus they assumed that if they found similar results that their assessment would indeed demonstrate construct validity. Based on their results, women were more able than men at preparing food and taking care of their appearance, hygiene, and house. Contrary to their findings, other researchers have found that there were no clinically detectable differences in ADL ability between men and women (Merritt & Fisher, 2002; Rice et al., 1998). Given the fact that there are discrepancies within the literature regarding gender differences in ADL ability, it seems suspect that Cry et al. (1994) would use this concept as the basis for construct validity. Furthermore, the authors never once introduced the idea of gender bias within the ILSS, for it is possible that the items within this assessment offer different levels of challenge to men and women, such that women may have an unfair advantage.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Assessment of Living Skills and Resources

The last assessment that will be discussed is that of the Assessment of Living Skills and Resources (ALSAR) (Williams et al., 1991). The developers of the ALSAR took a seemingly unique approach with their assessment, for they not only chose to assess the ability to *accomplish* specific community living tasks (skill), they also chose to investigate the available resources that individuals have at their disposal. Additionally, a risk score is calculated by considering both the individual's skills and resources. The specific community living tasks within the assessment include telephoning, reading, leisure activity, medication management, money management, transportation, shopping, meal preparation, laundering, housekeeping, and home maintenance. The authors define skill as "accomplishment of a task by the patient" (Williams et al., 1991, p. 85), which is scored using a 3-point rating scale. A score of zero indicates independent and consistent performance, which includes independently taking responsibility for the procurement of the task (e.g., arranging for another to perform the task). A score of 1 indicates partial performance or procurement, and a score of 2 indicates that the client does not perform nor take responsibility for the procurement of the task.

As noted above, within each task performance, a resource score is also assigned to each of the 11 tasks. Resource is defined as "a support for task accomplishment extrinsic to the patient" (Williams et al., 1991, p. 85). Resources may be technical or human, formal or informal, and include entities such as persons, equipment, services, and/or agencies. Within the ALSAR, resource level is also scored on a 3-point scale, with a score of zero indicating that resources are adequate to consistently accomplish the task (e.g., full-time caregiver prepares all of the meals, bus stop one block from the home). If

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

the resources only partially support the accomplishment of the task, and/or if the resource is inconsistent, unreliable, or unstable, then a score of 1 is assigned. Additionally, if a caregiver is experiencing strain/stress in response to providing the necessary care, then a score of 1 would be assigned. Lastly, if the resources are insufficient for task accomplishment, if the available resources are not being used, if the utilization of the resource presents a safety risk, and/or if loss of the resource is imminent, then a score of 2 is assigned.

In addition to rating the client's skill level and resource level, the examiner also assigns an overall rating of the client's risk within each task. Risk within the ALSAR is determined by assessing the client's skills and resources. The risk score ranges from zero to 4, and is obtained by summing the skill and resource scores for each task. A score of zero or 1 represents a low risk, a score of a 2 indicates moderate risk, and a score of 3 or 4 indicates high risk. The risk scores within each task are aimed to assist the healthcare professional in targeting and prioritizing goals and developing the most appropriate plan of care. After all of the risk scores have been calculated, they are summed to reveal the R-score, which is a "predictor of negative outcomes such as death, relocation to a more structured setting, or a change in supportive living arrangements from informal supports to more formal ones" (Williams et al., 1991, p. 88).

With an initial sample of elderly veterans, the researchers found that the ALSAR demonstrated adequate inter-rater reliability (86% agreement for the skill items and 95% agreement for the resource items) and internal consistency among the scale items (Cronbach's alpha = 0.91) (Williams et al., 1991). Additionally, the researchers concluded that the ALSAR has good predictive validity, in that the R-score was

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

significantly associated with negative outcomes during the 6 months following the evaluation, including nursing home placement (t = 1.64, $p \leq .05$), hospitalization (t = 1.76, $p \leq .05$), and death (t = 1.94, $p \leq .05$). Thus, it is thought that the ALSAR R-score may be useful in "alerting staff to the need for more extensive support systems such as more formal caregiver support or a move to community-based care with more services" (Williams et al., 1991, p. 90).

Overall, the ALSAR is an innovative approach to the assessment of living skills, because the individual is not "penalized" for arranging the procurement of daily living tasks. This concept is in line with a current philosophical trend in the concept of independent living, in that there is a shift from valuing a strictly independent community to a valuing a mutually dependent community (Townsend & Ryan, 1991). Given this stance, "independent living means the ability to direct, manage, or control the accomplishment of 'community living' through one's own actions or through directing the actions of others" (Townsend & Ryan, p. 53-54). Thus, within this paradigm, it is thought that the assessment of independence should not be determined by independent performance, but rather through personal control of decision-making and mutual dependence. However, given this stance, it becomes imperative that one *also* assess the available resources, for just because an individual has the skill to independently arrange for another to perform a task (or independently utilize public transportation, etc.), it does not mean that the resources are available to support this skill.

Additionally, within the initial phases of rehabilitation, it seems that one would also need to assess the client's *actual* ability to independently and skillfully perform everyday tasks. With this information, the rehabilitation team, client, and caregivers can

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

then determine the extent and type of resources that the client will likely need once he/she returns to the community. For example, when working with a client with a newly acquired injury/disability, healthcare providers must strive to build self-advocacy skills and assist individuals and their families in the determination of what resources are needed, wanted, and/or available (e.g., assistive technology, caregiver assistance, adaptive equipment, home modifications). If it is unknown what the individual can/cannot actually do and what specific skills he/she possesses, such activities within the treatment plan may be difficult. Thus, there appears to be pros and cons to giving "allowances" for the procurement of daily life tasks in the determination of independence within the community.

Overall, it is thought that the determination of risk, based on one's ability to perform and/or procure daily tasks and the available resources, is a valuable asset within client-centered assessment. If healthcare providers can utilize this information to prevent negative outcomes (e.g., rehospitalization, admission to a nursing home), perhaps more comprehensive plans for services and treatment could be made to prevent such occurrences.

Summary of Independent Living Assessments

Many important insights have been gained through investigating the numerous assessments of community living skills. First and foremost, it appears that the existence of a reliable, valid, sensitive, comprehensive assessment of one's ability to live independently within the community is either nonexistent or inaccessible via a literature search using of some of the popular data bases (e.g., Web of Science, Medline). This discovery is in line with other researcher's findings (Dickerson, 1997; Townsend &

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Ryan, 1991), for within their investigations, they also failed to find adequate assessments of global function within the community.

Within her investigation of several assessments of community function of individuals with mental illness, Dickerson (1997) stated "it is not surprising that all of the selected instruments contain at least some limitations . . . . It is unlikely that any one instrument can meet the demands for outcomes data across various settings." (p. 901). In fact, she stated that more effort should be directed toward sharing outcomes across settings and disciplines as this may be a more valid means of assessing community functioning. Likewise, Townsend and Ryan (1991) commented "in reviewing standardized, assessment tools before and after the study, I found none capable of addressing the holistic concept of 'independence in community living'" (p. 53). Given that the skills that make up independent community living are dynamic, complex, and intertwined, and given that every individual and every environment is unique and complex, the development of a single assessment that accurately depicts such skills is a very difficult task.

Secondly, the assessments reviewed above cannot be considered gold standard assessments with regard to determining level of dependence. The only assessment that has been utilized to predict level of care needed is the SLICLS. Researchers have documented that mean total SLICLS scores differed significantly between groups of individuals from three different residential settings (independent apartment, residential care facility, and nursing home) ($F = 32.9$, $p < .0001$) and the SLICLS scores significantly correlated with the social worker's estimate of the best level of placement ($r = .65$, $p < .001$) (Fitz & Evenson, 1995). Furthermore, a SLICLS-C score of 85

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

(sensitivity = 0.90, specificity = 0.81) was chosen as the cut point for independent versus semi-independent and a score of 66 (sensitivity = 0.91, specificity = 0.55) was chosen as the cut point for semi-independent versus dependent for a group of Chinese adults with schizophrenia (Au et al., 2005). Although the SLICLS has shown some initial promise in determining level of dependence, the validity of this tool has not been proven across a wide variety of clients. Furthermore, the documented accuracy of the predictions is inconsistent, as evidenced by the large standard deviations in the Fitz and Evenson (1995) study and the decreased accuracy of categorization with lower functioning clients in the Au et al. (2005) study.

Global Judgment of Functional Ability

The search for a gold standard assessment of independent living was not found, and thus alternative options must be considered. Several authors have noted the inherent problems of utilizing the client's current living environment and/or discharge setting as a gold standard for investigating the amount of assistance and/or level of care needed (Fitz & Evenson, 1995). In their investigation of the St. Louis Inventory of Community Living Skills (SLICLS), the mean scores of individuals residing in various settings had very large standard deviations, indicating that within seemingly similar settings there is a wide variation in the amount of assistance that individuals need. Likewise, an investigation of the Community Living Assessment Scale (CLAS), the authors found that there was more variation within the different settings (independent rooming house, family home, community residence, and nursing home) than there was between them (Willer & Gustaferro, 1989). Thus, it is possible that the "the label used to identify a setting does

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

not necessarily indicate the level of independence experienced by the individual living in that setting" (Willer & Gustaferro, 1989, p. 273).

Another option would be to utilize caregiver judgment about how much assistance individuals need. One obvious problem with this option is that many individuals do not have caregivers. For example, many older adults who are "living independently" within the community may be in need of some assistance to ensure safe and effective performance of all of the required independent living tasks (e.g., money management, heavy housework, yard care). Such individuals, although in need of assistance, may not be getting the assistance that is needed, and thus, may not have family members or other proxies who could accurately report their actual level of independence/need for assistance.

Within investigations of the utilization of caregiver judgment, authors have noted that caregivers have a tendency to declare higher rates of dependency (Cotter et al., 2002; Santos-Eggimann et al., 1998; Rubenstein et al., 1984). Furthermore, some also note that there may be a tendency of clients to underestimate levels of dependency (Santos-Eggimann et al.; Rubenstein et al.). Thus, it appears that the utilization caregiver and/or self-report questionnaires may not offer the most accurate rating of level of independence.

Clinical judgment, on the other hand, has been shown to have some promise in rating global level of independence. Global assessments of function, utilizing clinical judgment, offer quick and reliable ratings. However, the use of fewer categories to describe function may lead to reduced sensitivity (Domerick, Edwards, & Diringer, 2003). Regardless, there is supporting evidence for the use of global ratings, based on

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

clinical judgment. For example, within a study that examined the accuracy of medical doctors in rating global functional status, it was noted that the doctors were able to identify the need for assistance in most areas of basic ADL (feeding, bathing, dressing, and toileting) with sensitivity ranging from 81% to 92% (Pinholt et al., 1987). In the area of continence, the doctors displayed less ability to identify the need for assistance (sensitivity = 65%) and even less ability to correctly identify mental status impairments (sensitivity = 32%). It stands to reason that if medical doctors are able to identify the need for assistance in ADL (within a reasonable range of error), that perhaps occupational therapists, who are specialists in the assessment of ADL ability, would likely be even better judges of global functional status.

Another interesting study investigated the clinical judgment of occupational therapists in assessing participants' functional ability (Rogers et al., 2003). In this study, the occupation therapists were only given written information about the participants' body functions (e.g., sensory, cognitive, motor, affective abilities). With this limited amount of information, the therapists did not display the ability to accurately judge the participants' functional abilities, and quite often overestimated disability. The authors noted that the presentation of data related only to body functions could have led the therapists to exaggerate actual disability, or it is possible that impairments did not necessarily aid in the prediction of daily life function. Interestingly, the caregivers (who had directly observed daily life task performance of the participants) were able to more accurately judge disability than were the trained clinicians who had information only about impairments. It thus stands to reason that the clinical judgment of occupational therapists, regarding the assessment of global functional ability, would likely be more

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

accurate if they directly observed and assessed actual performance of daily life tasks (vs. the assessment of impairments). Such findings are in line with current trends within the field of occupational therapy, for occupational therapists are being called upon to focus their practice on occupational performance rather than disease and body function (Fisher, 1998; Neistadt & Seymour, 1995; West, 1984), and many authors have noted the failure of tests of impairment to predict disability (DeBettignies & Mahurin, 1989, Dickerson, 1997, Rice et al., 1998; Lindén et al., 2005; Searight & Golderberg, 1991; Winograd, 1984;)

In their study of clients with dementia, Hartman and colleagues (1999) used occupational therapists' clinical judgment via utilizing the AMPS global ratings of functional level. The AMPS global functional ratings are based on the clinical judgment of the occupational therapist as they assess their clients' abilities to live independently within the community. As outlined within the specific criteria in the AMPS manual (Fisher, 2006a, 2006b) the therapist must consider the client's ability to perform ***all*** of the necessary tasks and activities that are required of independent living within the community. The AMPS global functional levels are not based solely on the AMPS task observations. The ratings are based on ***all*** the information that the therapist has gathered about their clients, and thus, multiple sources can be utilized to come up with the most accurate rating. For example, the clinician can utilize information gathered through (a) interviewing the client and/or caregivers, (b) gathering information from other healthcare providers (e.g., case managers, nurses, physical therapists, speech therapists, doctors), (c) observing and assessing the client's actual ADL ability, and/or (d) assessing the available resources as well as the client's body functions. Given the fact that an AMPS

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

observation is conducted in conjunction with determining the client's global functional level, the therapist has direct knowledge regarding the client's ability to perform daily life tasks. It stands to reason, therefore, that their clinical judgment would be much more accurate than if they had simply used assessments of impairments. Furthermore, this approach is in line with current trends, for it is thought the utilization of multiple sources may be the most accurate means of determining functional independence (Dickerson, 1997).

The AMPS global functional levels are used to rate the client as either (a) able to live independently in the community, (b) needing minimal assistance to live in the community, or (c) needing moderate to maximal assistance to live in the community (Fisher, 2006a, 2006b) (Table 1). Obviously, the most ideal and accurate assessment of one's ability to live within the community would include direct observation of the client's ability to perform ***all*** the necessary tasks and activities necessary for independent living (Dever & Knapczyk, 1988), however, this is not clinically practical. Although the 3-point rating scale of the AMPS functional levels may have limited sensitivity, within this study, the occupational therapists' clinical judgment of global function was thought to be the most valid and practical option available.

Receiver Operating Characteristic Curves

In this section, I would like to take the time to discuss receiver operating characteristic (ROC) curves, a statistical method of determining the most accurate cutoff points for categorizing individuals into the appropriate levels of dependence. More specifically, I will discuss ROC curves from a general perspective and then I will place the utilization of ROC curves within the context of the current project. As an

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

introduction, ROC curves have traditionally been used within the medical profession as a means of determining the most accurate diagnostic clinical markers. However, researchers within the fields of rehabilitation, education, healthcare, public safety, and engineering have now begun utilizing ROC curves within their investigations about optimal decisions and diagnostic accuracy (Foster et al., 2004; Studenski, Wallace, Duncan, Rymer, Lai, 2001; van Bennekom, Jelles, Lnkhorst, & Bouter, 1996; Swets, Dawes, & Monahan, 2000).

In the most general sense, the ROC curve is a graphical depiction of the accuracy of a diagnostic test, depicting the rate of true positives (sensitivity) plotted against the rate of false positives (1 – specificity) (Appendix C and Figure 1) for any given decision threshold or cut point of a diagnostic test (Biagini, Kreig, Pinkerton, & Hamilton, 2001; Egan, 1975; Zhou, Obuchowski, & McClish, 2002). Thus, for each decision threshold, there is a corresponding prediction of the accuracy of the diagnostic test (true positive rate and false positive rate). The accuracy of the test is an indicator of the test's ability to discriminate between those with and those without the condition in question. Although there are many different decision thresholds, only two possible diagnostic decisions can be made; either the person has the condition or the person does not have the condition.

When examining the relationship between sensitivity and specificity, as the sensitivity increases, there is a resulting increase in the false positive rate (specificity decreases) (Figures 2 and 3). One unique aspect of the ROC curve is that estimates of the various outcomes (true positive rate, true negative rate, false positive rate, and false negative rate) are displayed for each cut point value. Thus, the curve allows the researcher to use previous probabilities of the conditions, in conjunction with the costs

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

and benefits of correct and incorrect decisions, to determine the best cut point for a given test (Zhou et al., 2002; Zweig & Campbell, 1993;).

The statistics generated from the ROC curve analysis have several benefits over traditional measures of accuracy. For example, the probability of a correct test result, the odds ratio, and Youden's index are all based on (and bound to) one decision threshold (cut point), all consider false negatives and false positives to be equally undesirable, and the former is dependent on the prevalence of the condition in the sample (Zou et al., 2002). Other advantages of the ROC curve over isolated measurements of sensitivity/specificity include (a) visual representation of accuracy, (b) selection of a specific decision threshold is not required, (c) estimates are not dependant on the scale of the test results (depends only on the ranks of the observations), (d) sample selection need not be representative of the prevalence of the condition, and (e) direct visual comparison of two or more tests on a common set of scales (Zou et al., 2002; Zweig & Campell, 1993).

The area under the ROC curve is used as a global indicator of the accuracy of a test and can range from 0.5 to 1.0, with an area of 1.0 indicating a perfectly accurate test. A test with perfect accuracy (sensitivity = 1.0) would produce a curve that reaches the upper left quadrant of the graph (Figure 1). In contrast, a test that is a purely random predictor (an equal rate of correct and incorrect decisions) would yield an area index of 0.50, and would be depicted by a straight diagonal line going from the origin to the top right corner of the graph (Figure 1). The area index is the most common global measure of the accuracy of a diagnostic tool (Zou et al., 2002; Zweig & Campell, 1993), and is used to depict the total error that may occur when using the assessment (the higher the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

area under the curve, the less error). The area index is not used to determine the type or amount of error at each specific cut-point.

The ROC curve area has several definitions and interpretations including, (a) the average value of sensitivity for the possible values of specificity, (b) the average value of specificity for all possible values of sensitivity, and (c) the probability that a randomly selected person with the condition has a test result that indicates that there is more suspicion of having the condition than one who does not have the condition (Zou et al., 2002). The last definition, expressed in simplified terms, means that if the area under the curve was 0.89, then a person with the condition would be 89% more likely to have a positive result than someone without the condition. When assumptions about the distribution are not made (e.g., nonparametric method), the area under the ROC curve is readily recognized as the Mann-Whitney version of the nonparametric two-sample statistic (Zhou et al., 2002; Zweig & Campbell 1993).

The area under the curve is a less precise index of the accuracy of a test, for the accuracy is condensed down into a single index, and thus, there is an inherent loss of information (Zweig & Campell, 1993). Due to the condensed nature of this measure, when considering the area under the ROC curve, it is desirable to also examine the plot itself. For example, two different tests may have the same total area (or different areas), but upon examining the clinically relevant region of the curve (e.g., the point on the curve that corresponds with the decision threshold), one may find that one test offers more desirable true positive and false positive rates (Zweig & Campbell, 1993; Zhou et al., 2002).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Because a "good" area index inherently depends on the disorder and the clinical use of the test, specific guidelines regarding desired or critical area values have not been globally accepted (Zou et al., 2002). However, rough guidelines for determining the discriminating value of a test via examining the area under the curve values are as follows: fail = 0.50 to 0.60, poor = 0.60 to 0.70, fair = 0.70 to 0.80, good = 0.80 to 0.90, and excellent = 0.90 to 1.00 (Perneczky et al., 2006; Tape, n.d.).

Area values are typically utilized early in the analysis of a diagnostic test, for the researcher must, first and foremost, determine if the test has merit or discriminating power. If the test cannot be used to discriminate between clinically relevant classes of individuals, then there is little value in continuing the investigation of the test's clinical use (Zweig & Campbell, 1993). If a diagnostic tool has merit, the next step is to critically investigate the clinical usefulness of the test, and then decide upon a decision threshold. When investigating the clinical usefulness of a tool, the researcher must examine the practical value of the test and the information gained from using the test. For example, a test may have good discriminating power, yet be of little practical use for patient care due to the cost of using the test, the invasive nature of the test, the availability of the test, the availability of trained evaluators, or the relative cost of an incorrect diagnosis (Zweig & Campbell, 1993). The usefulness of the test involves many financial, clinical, personal, and societal factors and thus, when examining the value a diagnostic test, it is important to separate the inherent characteristics of the test (e.g., accuracy) from the practical usefulness of the test.

Another use of the ROC curve is to determine the optimal threshold or decision point. In ideal situations, the optimal decision threshold is an informed decision that

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

takes into account the prevalence of the condition in the target population and the relative costs or consequences of the test results (e.g., the cost of false positive and/or false negative decisions) (England, 1988; Greiner, Pfeiffer, & Smith, 2000). One means of determining the optimal cutoff value of a diagnostic test is the sensitivity at a fixed or predetermined false positive rate. This index allows researchers to identify the clinically desired false positive rate and investigate the accuracy of the test at that point on the curve (Zhou et al., 2002). If this statistic is utilized, the researcher predetermines the desired false positive rate based on the prevalence of the disorder and/or the consequences of an incorrect diagnosis (Zhou et al., 2002; Zweig & Campbell, 1993). Failing to predetermine the fixed false positive rate could lead to potentially biased results, as the researcher's beliefs about the test and the results may affect decisions (Hanley, 1989; Zhou et al., 2002). An advantage of using the sensitivity at a fixed false positive rated is that it is relatively simple and offers a clinically useful interpretation of the data. Disadvantages include (a) other studies may report sensitivity at different false positive rates, and thus it becomes more difficult to compare results; (b) published reports are not always clear as to whether the fixed false positive rate was predetermined; and (c) statistical reliability (e.g., more variation) is lower than that of the area under the ROC curve (Hanley, 1989; Zhou et al., 2002).

If the researcher is unable to determine a desired fixed false positive rate, and if the relative prevalence of the condition can be assumed to be 0.50 (e.g., 50% of the population has the condition), then one may determine the optimal cutoff point via selecting a point on the ROC curve that is closest to the upper left corner. This point on the curve corresponds to the cutoff threshold that maximizes both sensitivity and

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

specificity (Greiner et al., 2000). Although the practice of giving equal weights to sensitive and specificity facilitates comparison to other diagnostic tests, this method can lead to not utilizing the diagnostic test to the fullest potential within a given population/condition (Greiner et al., 2000).

If, on the other hand, the relative costs of the possible decisions are known, calculations can be made to determine the optimal point on the ROC curve that minimizes the overall cost of using the test (England, 1988; Greiner et al., 2000; Zhou et al., 2002) (the equations for determining the optimal cutoff measure are readily available in the references cited). However, relative costs may not be known or quantifiable. According to Greiner and colleagues (2000), "the challenge of this approach is that it requires the users to quantify the consequences of each possible test outcome, although outcomes are often thought of only qualitatively" (p. 31). Given the fact that (a) it is often difficult to quantify decision costs and benefits, (b) most prevalence rates are not at 0.50, and (c) treating sensitivity and specificity equally may not be warranted given the prevalence rates and relative costs, many defer to the more practical method of examining the clinically relevant portion of the ROC curve (the area of the curve with acceptable false positive and true positive rates) to determine the optimal threshold or cutoff measure (Pepe, 2003; Zhou et al., 2002).

To further investigate the concept of the clinically relevant area of the ROC curve, it is important to note that the costs and benefits can be viewed from the perspective of the patient, caregivers, insurance companies, and/or society. For example, if a false negative test result is obtained, and death results because medical services were not delivered, the cost of this type of error is large, especially if the available treatment could

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

have prevented death. In general, if the prevalence of the disorder is relatively high, or if the cost of a false negative result is greater than the cost of a false positive result (such as the case in this example), then a low decision threshold is desirable. In this instance, one would favor sensitivity, for one would want to maximize the true positive rate and minimize the false negative rate (Figure 2) (Biagini et al., 2001; Greiner et al., 2000; Zweig & Campbell, 1993).

In contrast, if a false positive test is obtained (the test result indicates that the person has the condition when in reality they do not have the condition), and the medical treatment is expensive and/or harmful to healthy patients (e.g., chemotherapy, surgery), the result could lead to higher financial burden for the patient and insurance company, and may even result in higher insurance rates for the general population. Additionally, the administration of the unneeded treatment could lead to unnecessary illness or harm to the person. In this situation, the cost of a false positive is relatively high. In general, if the prevalence of the condition is low and the cost of a false positive result is greater than the cost of a false negative result, a higher decision threshold is desirable. Thus, in this situation, one would want to favor specificity over sensitivity, as a lower false positive rate would be the most advantageous (Figure 2) (Biagini et al., 2001; Greiner et al., 2000; Zweig & Campbell, 1993).

Lastly, when generating and analyzing ROC curves, one must determine the type of data being analyzed (ordinal or interval), and if the data can be assumed to follow a normal distribution. Using continuous data offers an advantage, for with this type of data there is a smaller chance of generating a "tie." A tie occurs when a person within the condition group has the same test result as a person without the condition (Zweig &

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Campbell, 1993). When generating a curve, a tied sample introduces potential error, as there is no way of distinguishing the two individuals.

To deal with the issue of tied data, some researchers have used a parametric model for fitting a curve (a smooth curve is generated vs. a stepwise curve), such that the parameters are estimated based on the notion that the distribution follows a normal or predictable pattern (Zhou et a., 2002; Zweig & Campbell, 1993). If continuous data are utilized, however, Zweig and Campbell note that it does not make sense to collapse the data to fit a parametric model. Thus "for continuous data, the nonparametric ROC is preferred" (Zweig & Campbell, 1993, p. 567), for parametric assumptions are not utilized and unbiased estimates of sensitivity, specificity, and area under the curve can be obtained. One disadvantage of using the nonparametric approach is that, although tied data are much less likely, if ties occur the accuracy of the test may be underestimated (Zhou et al., 2002; Zweig & Campbell, 1993).

Decision Thresholds – Considerations within the Current Study

In the following section, I will discuss the clinically relevant area of the curve, as it pertains to the current study. When an assessment or test is used to classify individuals into different groups, there is always a potential for error/incorrect decisions. To help clarify the different types of error and how changing the cut-point impacts the error rates, the reader is encouraged to refer to Figures 2 and 3 during the following discussion. When using a test to classify individuals into one of two groups, the sample becomes split into those with the condition and those without the condition. Within the current study, the sample may be divided into those who are independent and those in need of assistance. In Figures 2 and 3, one can see that the sample that is independent in the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

community is on the left side of the ADL ability scale, while those in need of assistance are on the right side of the scale. The horizontal line running through figure depicts the cutoff measure; those with ADL ability measures above the cutoff measure would be coded as being independent (a positive test result) those below the cutoff measure would be coded as needing assistance (a negative test result).

In examining Figure 2, when a person in the independent sample is correctly coded as being independent, a true positive decision is achieved. In contrast, when a person in the independent sample is incorrectly coded as needing assistance (i.e., their ability measure falls below the cutoff measure), a false negative decision is made. Likewise, if a person in need of assistance has an ADL ability measure above the cutoff measure, he/she would be incorrectly coded as being independent (false positive decision). Lastly, if a person that is in need of assistance has an ADL ability measure below the cutoff, he/she will be correctly coded as needing assistance (true negative decision). If the cutoff measure is moved up or down, it will have an impact on the error rates. For example, if the decision threshold or cutoff measure is raised (i.e., create a decision threshold that corresponds to higher ADL ability), the result is a lower false positive rate at the cost of a higher false negative rate (Figures 2 and 3). In contrast, if the decision threshold is lowered, we trade a decreased false negative rate for a higher false positive rate. Thus, when creating a decision threshold or cutoff measure, one must decide upon a point where the opposing error rates are "balanced." Such a point depends on the clinical use of the tool and the characteristics of the population.

In the current research project, the AMPS was investigated to see if the resulting ADL motor and/or ADL process ability measures can discriminate between (a)

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

individuals who need assistance and those who do not need assistance, (b) individuals who are independent and those who need minimal assistance, and (c) individuals who need minimal assistance and those in need of moderate to maximal assistance. Within this investigation, the cost of a false positive decision will be considered greater than the costs of a false negative decision (Figure 4). To review, a false negative error occurs when the individual's test result indicate the need for assistance, when in fact the person is capable of living independently within the community. In contrast, a false positive result indicates that the person is able to live independently within the community, when in fact they need some assistance or support to safely and effectively live within the community. In support of my decision to consider a false positive decision to be more costly than a false negative, Dijkstra, Tiesinga, Plantinga, and Veltman (2006), in their investigation of the Care Dependency Scale, a nursing assessment of dependency needs for elderly clients, also chose cutoff values that minimized errors that ultimately lead to clients not receiving the care and support that they actually needed.

To further explore this issue, it is important to note that a false positive error not only results in the client returning home alone (or maintaining "independent" living) when there is a need for assistance, it may also result in the loss of receiving valuable intervention. Such circumstances could result in injury or hardship to the client and/or others within the household. For example, the client could injure him/herself, fail to prepare and/or eat an adequate diet, leave a burner on and start a fire, mismanage finances, cause an automobile accident, etc.

To elaborate further, one potential outcome of a false positive error is that the client will experience a fall. One can quickly see that the costs of such an occurrence are

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

fairly substantial to both the individual and society. Many authors have documented the potentially devastating results of falls within the aging community, including debilitating injuries, loss of independence, institutionalization, and death (Fortinsky et al., 2004; Haentjens, Autier, Barette, & Boonen, 2001; Weir & Culmer, 2004). In terms of financial costs, one study reported that for individuals with hip fractures, the mean initial hospitalization cost per person was $9534, and the mean cost above typical medical expenses for the year following hospitalization was $7300 (Haentjens et al., 2004). Others stated that in 1999, Medicare expenditures related to falls was in excess of $8 million, and costs related specifically to fractures were in excess of $5 billion (Bishop, Gilden, & Blom, 2002). Thus, if such potentially damaging events can be prevented, through additional therapy and/or through the provision of assistance, then steps should be taken to ensure that individuals are identified as needing such services.

In contrast, the costs associated with false negative errors are likely not as high as those associated with false positive errors. For example, if a client returns home with family assistance, home therapy, and/or home healthcare, continued assessment of the client's level of safety and dependence can lead to an eventual correct decision and the termination of services once the person has demonstrated safe and independent performance of the tasks required for independent living. The costs of providing in home assistance for a short period of time are likely much less than the costs that can occur as a result of a false negative error. To support this notion, within their study of the economic costs of hip fractures, home therapy accounted for the smallest percentage of the total costs (14%) (Haentjens et al., 2001). Thus, it stands to reason that within this study, a decision threshold that seeks to minimize the rate of false positives is logical.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Lastly, it is important to note that a negative test result (e.g., the person is below the cutoff measure) could be used to document (a) the need for and/or level of assistance required to live in the community, or (b) the need for further therapy or treatment so that the need for assistance can be minimized or eliminated. Such decisions ultimately fall on the hands of the occupational therapist and the rehabilitation/clinical team, for if a client demonstrates the potential for improved ADL ability, then continued therapy may be recommended. In contrast, if the client has reached a functional plateau or if the person does not want to continue therapy, then appropriate measures could be taken to ensure that the person has the assistance that is needed.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

CHAPTER 3

METHODS

Research Design

The research questions in this investigation include (a) Is there a significant association between ADL ability (motor and/or process) and global functional level?; (b) Can AMPS ADL motor and/or process ability be used to correctly categorize individuals who need assistance and those who do not need assistance to live in the community?; and (c) Can the ADL motor and/or ADL process ability measures be used to accurately determine level of dependence within the community (i.e., independent vs. minimal assist vs. moderate to maximal assistance)?; and (d) When using ADL motor and/or ADL process ability measures to determine level of community dependence, are there differences in area under the ROC curve estimates among the different diagnostic groups? Globally, these questions ask if the ADL motor and/or ADL process ability measures can be used to accurately categorize individuals into varying levels of community independence. To do this, I used ROC curves to determine the statistical probability of correctly categorizing individuals into the varying levels of independence through the utilization of AMPS ADL motor and ADL process ability measures.

Research questions that specifically ask about the reliability, validity, and/or the prognostic indicators of assessments typically lead researchers toward the utilization of quantitative methodologies (Bennett & Bennett, 2000; Sackett, Rosenberg, Gray, Haynes, & Richardson, 2000). As the current research questions (noted above) inquire about the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

validity of the AMPS ADL ability measures and require the generation of statistical probabilities, I utilized a quantitative methodology in this investigation. Clearly, qualitative methods would not be appropriate in answering the research questions, for such methods are used to help the researchers make sense of, or interpret, phenomena in terms of the meanings that people associate with them (Green & Britten, 1998).

For this project, I used a convenience sample of existing data where the AMPS ability measures and global functional level ratings were collected concurrently. Through using global ratings of functional level as the external criterion, I investigated the evidence of utilizing the AMPS ADL ability measures in the determination of level of community independence. More specifically, I conducted a retrospective, criterion-based validation study of the ADL motor and ADL process ability measures of the AMPS.

Participants and Site

The sample used in this study consisted of existing data, located in the AMPS Project International database in Fort Collins, Colorado. As described in more detail in the instrumentation section, the data was from participants who were assessed within typical occupational therapy settings (e.g., hospitals, client homes, rehabilitation clinics).

The participants potentially included all available people in the Assessment of Motor and Process Skills (AMPS) database who were evaluated between January 1, 1999 and January 1, 2005 and who met the following criteria: (a) were 16 years of age or older; (b) were not associated with rater scoring error as evidenced by artificially high ADL motor or ADL process ability measures; (c) were not scored by multiple raters as part of rater calibration co-scoring; and (d) had known gender, age, and global functional level ratings (i.e., independent, in need of minimal assist, or in need of moderate to

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

maximal assist). Participants of varying diagnoses (including well, healthy adults) and varying functional levels were included in the potential sample.

The sample of participants was chosen to reflect the wide range of clients that are typically assessed within settings were occupational therapy services are provided (e.g., private homes, nursing homes, acute-care hospitals, rehabilitation centers, community mental health centers). Additionally, sample participants were tested by calibrated AMPS raters from North America, Nordic countries, Europe, Australasia (Australia and New Zealand), Asia (Taiwan, Japan, Hong Kong, Republic of Korea, and the Philippines), Israel, South Africa, or South America. Through utilizing a diverse sample, it was believed that the clinical application of the research study findings would be more powerful, as occupational therapists would be able to apply the findings across the diverse client populations they serve. Furthermore, to examine the predictive value of the AMPS across all levels of community independence, I needed a sample that included individuals of varying ADL abilities from all three global functional level ratings. Lastly, a diverse sample also offered the potential to conduct additional secondary analyses to determine if the predictive value of the AMPS differed among different client populations.

Children below the age of 16 were excluded from the sample as the AMPS computer-scoring software (Three Star Press, 2005) has default functional level ratings that cannot exceed a predetermined level of independence. For example, a 5-year-old child cannot have a functional level rating of (a) independent in the community or (b) requires minimal assistance to live in the community (i.e., the only appropriate functional level rating would be that the child needs moderate to maximal assistance to live in the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

community) (Table 1). Likewise, a 10-year-old child can only have functional level ratings of either minimal assistance or moderate to maximal assistance required for community living. Lastly, calibration data were excluded from the sample as clients who are scored by multiple raters can lead to biased estimates of the client's ADL motor and ADL process ability measures.

For this study, 64,446 participants (58% of in the AMPS database) met the criteria for inclusion. Participants eliminated from consideration consisted of those who were less than 16 years of age (23% of the total sample), were associated with rater scoring error (10% of the total sample), or had incomplete demographic data (9% of the total sample). In the final sample, 10,458 were rated as being independent within the community, 23,679 were rated as needing minimal assistance, and 30,329 were rated as needing moderate to maximal assistance to live in the community. Demographic information for the participants is documented in Tables 15 and 16.

Instrumentation

Global Functional level

Each participant's global functional level was rated by trained and calibrated AMPS raters according to the specific criteria outlined within the AMPS manual (Fisher, 2006b) (Table 1). According to the procedures outlined in the AMPS manual, when determining the AMPS global functional level rating the therapist must consider the client's ability to perform ***all*** of the necessary tasks and activities that are required of independent living within the community (Fisher, 2006a, 2006b). Thus, the AMPS global functional levels are not solely based on the AMPS task observations. Rather, the rating is based on all of the information that the therapist has gathered about the client.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

As a result, the therapist must utilize multiple sources of information to determine the most accurate rating. For example, the clinician can utilize information gathered through (a) interviewing the client and/or caregivers, (b) gathering information from other healthcare providers (e.g., case managers, nurses, physical therapists, speech therapists, doctors), (c) observing and assessing the client's actual ADL ability (e.g., the AMPS observation, informal assessment of ADL), and/or (d) assessing available resources as well as the client's body functions. This approach is in line with current trends, for it is thought that the utilization of multiple sources may be the most accurate means of determining level of independence (Dickerson, 1997).

In considering all of the information gathered about the client, and using the specific criteria outlined in the AMPS manual (Fisher, 2006a, 2006b), the therapist used his/her clinical judgment to determine if the client is likely (a) able to live independently in the community, (b) in need of minimal assistance or supervision to live in the community, or (c) in need of moderate to maximal assistance to live in the community (Table 1). Global functional level ratings were determined prior to generating the client's ADL motor and ADL process ability measures. Thus, the AMPS rater was unaware of the client's specific ADL ability measures when the global functional level ratings were assigned.

Assessment of Motor and Process Skills (AMPS)

The AMPS was administered by trained and calibrated occupational therapy practitioners according to standardized procedures described in the AMPS manual (Fisher, 2006a, 2006b). Each AMPS evaluation consisted of several steps, the first being the occupational therapy interview. During this interview, the therapist inquired about

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

the types of tasks that the client performs on a daily basis and those tasks that present a challenge to the client. The therapist determined which personal and instrumental ADL tasks the client was interested in and familiar with, and the therapist then identified several (usually three to six) appropriately challenging tasks from among the 83 ADL tasks listed in the AMPS manual that were relevant, familiar, and discussed by the client. From the shortened list that was presented to the client as possible tasks to perform, the client chose at least two tasks, and the therapist and the client then agreed upon the specific task criteria for each chosen task. The client performed each task in a familiar environment and in his or her usual manner.

AMPS assessments were completed within typical occupational therapy settings, (e.g., client homes, hospitals, rehabilitation clinics). The AMPS can be validly used in various assessment environments, provided that the client (and therapist) has been fully familiarized with the set up of the environment, including the location of all needed tools and materials (Darragh, Sample, & Fisher, 1998; Nygård, Bernspång, Fisher, & Winblad, 1994; Park, Fisher, & Velozo, 1994; Robinson & Fisher, 1996). Familiarizing the client with the environment is part of the standardized procedures clearly documented in the AMPS manual (Fisher, 2006a, 2006b).

After completion of each AMPS task performance, the rater scored each of the 16 ADL motor and 20 ADL process skill items according to the criteria in the AMPS manual (Fisher, 2006b) using a 4-point ordinal scale. A score of 1 indicates deficient performance and/or task breakdown, a score of 2 indicates ineffective performance, a score of 3 indicates questionable performance, and a score of 4 indicates competent task performance. Thus, each participant had at least two sets of ADL motor scores and two

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

sets of ADL process scores, one set for each task performed. When the data were analyzed, both sets of scores were considered simultaneously, such that one ADL motor ability measure and one ADL process ability measure were generated. By using the results of two (or more) tasks, greater reliability of the estimates was achieved.

The AMPS is a criterion-referenced measure of ADL ability. Thus, the clients were scored according to an external criterion of competence. Within the AMPS, the ADL motor and ADL process skill scores were not based on norms; the scores were based on the criterion of no observable increased effort, decreased efficiency, safety risk, and/or need for assistance (i.e., competence) (Fisher, 2006a). Therefore, a score of 4 represented safe, independent, efficient, and effortless performance of the ADL motor or ADL process skill.

Although much of this information was discussed earlier within Chapter 2, a review of the reliability and validity of the AMPS also has been reported here. The AMPS has been shown to consist of two unidimensional scales of personal and domestic ADL ability (Fisher 1993, 1994, 2003a). Previous research studies have supported the reliability and validity of the AMPS ability measures across age groups (Dickerson & Fisher, 1993; Hayase et al., 2004), between men and women (Duran & Fisher, 1996; Merritt & Fisher, 2004), with a variety of diagnoses (Cooke et al., 2000; Doble, et al., 1994; Doble et al., 1997; Girard et al., 1999; Hartman et al., 1999; Kottorp et al., 2003; Kottorp et al., 1995; McNulty & Fisher, 2001; Oakley et al., 2003; Pan & Fisher, 1994; Robinson & Fisher, 1999; Robinson & Fisher, 1996; Sellers et al., 2001), and across cultural/ethnic groups and world regions (Bernspång & Fisher, 1995b; Dickerson & Fisher, 1995; Fisher, Liu, Velozo, & Pan, 1992; Goldman & Fisher, 1997; Goto et al.,

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

1996; Magalhães et al., 1996; Stauffer et al., 2000). Lastly, the AMPS has also been shown to be as a sensitive outcome measure (Graff, Vernooij-Dassen, Hoefnagels, Decker, & de Witte, 2003; Hariz, Bergenheim, Hariz, Lindberg, 1998; Kinnman, Andersson, Wetterquist, Kinnman, Andersson, 2000; Kottrop, Hällgren, Bernspång, & Fisher, 2003; Oakley, Khin, Parks, Bauer, & Sunderland, 2002; Oakley & Sunderland, 1997; Tham, Ginsburg, Fisher, & Tegnér, 2001).

Procedure/Data Analysis

Per standard protocol, prior to beginning this research project, I received approval from the Regulatory Compliance Office at Colorado State University. The data was submitted to the AMPS International Project according to the following procedures. After scoring each of the ADL motor and ADL process skill items, each rater entered their client's data into their personal, password protected application of the AMPS computer-software (Three Star Press, 2005). The raw ADL motor and ADL process scores and client demographic information (i.e., global functional level, gender, age, world region, and diagnosis) were then exported and sent to the AMPS Project International office in Fort Collins, Colorado. During the data export process, the AMPS computer-software strips all identifiable information (e.g., client name, unique client identification numbers) from the data, thus the client's identity was strictly maintained within each rater's password protected, personal application of the AMPS computer-scoring software (Three Star Press, 2005).

Upon receiving the data at the AMPS Project International office, the data were imported into the database. A specialized many-faceted Rasch (MFR) analysis program FACETS (Linacre, 2005) was used to convert the raw ordinal ADL skill item scores into

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

linear ADL motor and ADL process ability measures (Fisher, 1993, 1997; Linacre, 1993). Such analyses adjust the final ADL ability estimates to account for task challenge, skill item difficulty, and the severity of the rater. Participants were then selected from the database, according to the criteria noted above. Participant demographic information and ADL motor and ADL process ability measures were imported into the SPSS version 12.0 program for Windows.

To evaluate the association between ADL ability (motor and process) and global functional level, Spearman's rho correlations were calculated using SPSS version 12.0 for Windows. The resulting correlations were evaluated for (a) statistically significant associations ($p < .05$) and (b) effect size (r). Effect sizes were critiqued based on Cohen's guidelines (1988) where (a) $r = .1$ indicates a small effect, (b) $r = .3$ indicates a medium effect, and (c) $r = .5$ indicates a large effect.

Receiver operating characteristic (ROC) curves were then generated to evaluate the accuracy of using AMPS ADL ability measures to predict level of dependence in the community. Using SPSS Version 12.0 for Windows, ROC curves for both ADL motor and ADL process abilities were generated for the each of the following classification conditions (a) independent versus in need of assistance, and (b) other (i.e., independent and minimal assistance) versus in need of moderate to maximal assistance. As the AMPS ability measures are linear measures of ability, a nonparametric model was utilized to generate the ROC curves so that unbiased estimates could be generated (Zweig & Campbell, 1993).

The logic behind the classification conditions was the desire to (a) analyze the accuracy of the current ADL motor and ADL process cutoff measures demarking

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

independence versus in need of assistance, (b) determine if the accuracy of the cutoff measures demarking independent versus in need of assistance could be improved upon, and (c) determine the cutoff measure that could be used to identify those who were in need of moderate to maximal assistance. If the ADL motor and ADL process scales on the AMPS Graphic Report each had a cutoff measure demarking those who are independent and another demarking those in need of moderate to maximal assistance, then it would stand to reason that individuals falling between these two cutoff measures would likely need minimal assistance to live in the community. Analyzing the data in this manner allowed me to include the entire sample in each of the ROC curve analyses.

To generate ROC curves for the first classification condition of independent versus in need of assistance, global functional level was recoded into a dichotomous variable. Thus, functional level ratings of minimal assistance and moderate/maximal assistance were recoded into one group (n = 54,008) and those with functional level ratings of independent formed the second group (n = 10,458). The new dichotomous variable was then used to generate two ROC curves, one with ADL motor ability as the test variable and one with ADL process ability as the test variable.

The ROC curves for independent versus in need of assistance were first examined to globally determine if the ADL motor and ADL process ability measures have merit in identifying clients who are independent. Rough guidelines for determining the discriminating value of a test via examining the area under the curve values are as follows: fail = 0.50 to 0.60, poor = 0.60 to 0.70, fair = 0.70 to 0.80, good = 0.80 to 0.90, and excellent = 0.90 to 1.00 (Perneczky et al., 2006; Tape, n.d.).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Next, ROC curves were generated to determine if AMPS ability measures can be used to accurately categorize individuals as needing moderate to maximal assistance. To generate ROC curves for the second classification condition of other (independent and minimal assist) versus in need of moderate to maximal assistance, global functional level was once again recoded into a dichotomous variable. Thus, functional level ratings of independent and minimal assistance were recoded into one group (n = 34,137) and those with a functional level rating of moderate to maximal assistance formed the second group (n = 30,329). The new dichotomous variable was then used to generate two ROC curves, one with ADL motor ability as the test variable and one with ADL process ability as the test variable. Once again, the ADL motor and ADL process scales were globally evaluated by first examining the area under each ROC curve to determine if the scales can be used to identify level of independence.

Although specific guidelines or recommendations of desired levels of sensitivity and specificity were not found within the literature, other researchers have determined optimal cut points by choosing test values that minimized error and optimized correct decisions (i.e., equal weights were given to sensitivity and specificity) (Papaioannou et al., 2004; Perneczky et al., 2006; Uyttenboogaart, Stewart, Vroomen, Keyser, & Luijckx, 2006). In the current study, I chose to take an approach that was a bit more specific. Thus, upon verification that the ADL motor and/or ADL process scales had merit in predicting independence in the community and/or level of dependence (i.e., area under the curve $\geq$ 0.70), decision threshold values were determined by examining the ROC curve and the corresponding values for sensitivity and specificity at the clinically relevant area of the curve. In this study, the clinically relevant area of the curve corresponded to

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

the point on the curve where false positive rates were minimized, without resulting in large false negative rates. The rationale for favoring small false positive rates was discussed in detail in Chapter 2, and supported by methods utilized by Dijkstra and colleagues (2005).

Within the ROC curve analyses, I did not specifically assess the accuracy of correctly categorizing individuals who need minimal assistance to live in the community. Thus, to continue evaluating the accuracy of the new cutoff measures, I first generated new (predicted) functional level ratings based on the chosen cutoff measures, and then examined how the new functional level estimates compared to the actual ratings given by the occupational therapists. Lastly, to see if the estimates could be improved upon, I investigated the accuracy of the cutoff measures when matched decisions occurred (e.g., both ADL motor and ADL process abilities were either above or below the cutoff measure).

Lastly, ROC curves were generated for individuals within specific diagnostic categories. More specifically, for the two classification conditions listed above, I generated separate ROC curves from the data of clients within each diagnostic category. For this analysis, the area under the ROC curves was examined to globally determine the relative predictive validity across the different diagnostic categories.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

CHAPTER 4

RESULTS

Correlation analysis revealed that global functional level has a statistically significant association with ADL motor ability ($r = .35$, $p < .01$) and ADL process ability ($r = .43$, $p < .01$). According to Cohen's criteria (1988), ADL motor and ADL process ability measures have a medium effect with regard to global functional level.

Decision: Independent versus the Need for Assistance

The ROC curve for examining the utility of using ADL motor ability measures to determine community independence is shown in Figure 5. The area under the curve for this analysis was 0.78, indicating good discriminating value. Examination of the ROC curve revealed that the sensitivity (i.e., true positive rate or the rate of those correctly identified as being independent within the community) at the current ADL motor cutoff measure of 2.00 logits was 0.40, and specificity (i.e., true negative rate or the rate of those correctly identified as needing assistance to live in the community) was 0.87. Upon examination of the ROC curve, the ideal cutoff measure for ADL motor ability was determined to be 1.50 logits. Sensitivity at this cutoff measure was estimated to be 0.67 and specificity was estimated to be 0.72. Documented within Table 17 are the estimated rates of sensitivity and specificity. Refer to Figures 2 and 4 for a graphical depiction of the concepts of sensitivity and specificity.

When using ADL process ability measures to determine community independence, the area under the ROC curve was 0.84, indicating excellent

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

discriminating value (Figure 5). Examination of the ROC curve revealed that the sensitivity at the current ADL process cutoff measure of 1.00 logits was 0.81 and specificity was 0.70 (Table 17). Upon examination of the ROC curve, the ideal ADL process cutoff measure was determined to be 1.10 logits. Sensitivity at this cutoff measure was estimated to be 0.78 and specificity was estimated to be 0.73.

Upon determining the optimal cutoff measures for determining community independence, a new variable of "predicted independence" was created and the data were coded to represent the predicted functional level ratings for each person. More specifically, if the ADL motor ability and/or the ADL process ability were below the cutoff measures of 1.5 and 1.1 logits, respectively, then the data were coded to indicate a predicted need for assistance. Likewise, if the ADL motor and/or ADL process ability were at or above the new cutoff measures, then the data were coded to indicate predicted independence in the community. To continue the analysis of the new cutoff measures, I selected only those individuals with predicted ADL motor and ADL process decisions that matched. That is, only participants who had both ADL motor and ADL process ability measures above the new cutoff points or both ADL motor and ADL process ability measures below the new cutoff points were selected. Sixty-six percent of the sample had matched predicted decision points. When comparing the predicted ratings to the occupational therapists' original functional level ratings, 83% of the independent sample (n = 6,898) was accurately predicted and 83% of the sample in need of assistance (n = 35,516) was accurately predicted.

Lastly, we noted that of all of the subjects with ADL motor ability measures below the cutoff of 1.50 logits, 92% were correctly coded as needing assistance (n =

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

42,221). Of all of the subjects with ADL process ability measures below the cutoff of 1.10 logits, 94% were correctly coded as needing assistance (*n* = 43,395). Finally, of all of the subjects with both ADL motor and ADL process ability measures below the respective cutoff measures, 96% were correctly coded as needing assistance (*n* = 31,782).

Decision: Other versus Moderate to Maximal Assistance

A ROC curve was generated to examine the utility of using ADL ability measures to identify individuals who are in need of moderate to maximal assistance to live in the community (Figure 6). When using ADL motor or ADL process ability to identify those in need of moderate to maximal assistance, the areas under the ROC curves were 0.74 (good discrimination power) and 0.82 (excellent discrimination power) respectively (Table 17). The chosen ADL motor cutoff measure, based on this ROC curve analysis, was 1.00 logits, which corresponded to an estimated sensitivity of 0.70 and specificity of 0.66. The chosen ADL process cutoff measure was 0.70 logit, corresponding to an estimated sensitivity of 0.79 and specificity of 0.69.

Upon determining the optimal cutoff measures for identifying those in need of moderate to maximal assistance, a new variable was created and the data were coded to represent the new predictive functional level ratings for each person. That is, if the ADL motor ability and/or the ADL process ability were below the cutoff measures of 1.00 and 0.70 logits, respectively, then the data were coded to indicate the predicted need for moderate to maximal assistance. Likewise, if the ADL motor and/or ADL process ability were at or above the new cutoff measures, then the data were coded to indicate the prediction of “other” (independent and needing minimal assistance). To continue the analysis of the new cutoff measures, only those individuals with predicted ADL motor

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

and ADL process decisions that matched were selected. That is, I only selected participants who had both ADL motor and ADL process ability measures above the new cutoff points or both ADL motor and ADL process ability measures below the new cutoff points. Sixty-six percent of the sample had matched predicted decision points. When comparing the predicted ratings to the occupational therapists' original functional level ratings, 88% of the "other" sample (n = 23,097) was accurately predicted and 77% of the sample in need of moderate to maximal assistance (n = 17,992) was accurately predicted.

Lastly, we noted that of all of the subjects with ADL motor ability measures below the cutoff at 1.00 logits, 67% were correctly coded as needing moderate to maximal assistance (n = 29,161). Of all of the subjects with ADL process ability measures below the cutoff of 0.70 logits, 74% were correctly coded as needing moderate to maximal assistance (n = 27,776). Finally, of all of the subjects with both ADL motor and ADL process ability measures below the lower cutoff measures, 83% were correctly coded as needing assistance (n = 16,775).

Accuracy of Multiple Cutoff Measures

Next, multiple ADL motor cutoff measures were used to code individuals into one of three predicted categories (independent, in need of minimal assistance, or in need of moderate/maximal assistance). More specifically, the predicted functional level ratings were determined using the following criteria. Individuals with ADL motor ability measures $\geq$ 1.50 logits were coded as being independent, individuals with ADL motor ability measures that were $<$ 1.50 logits and $\geq$ 1.00 logits were coded as needing minimal assistance, and those with ADL motor ability measures $<$ 1.00 logits were coded as needing moderate/maximal assistance. When compared to the actual/original functional

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

level codes, 68% of the independent sample was correctly predicted (Figure 7), 25% of the sample in need of minimal assistance was correctly predicted (Figure 8), and 64% of the sample in need of moderate/maximal assistance was correctly predicted (Figure 9).

Multiple ADL process cutoff measures were then used to code the data into one of three predicted functional level categories. More specifically, the predicted functional level ratings were determined using the following criteria. Individuals with ADL process ability measures $\geq$ 1.10 logits were coded as being independent, individuals with ADL process ability measures that were < 1.10 logits and $\geq$ 0.70 logits were coded as needing minimal assistance, and those with ADL process ability measures < 0.70 logits were coded as needing moderate/maximal assistance. When compared to the actual/original functional levels, 75% of the independent sample was correctly predicted (Figure 7), 33% of the sample in need of minimal assistance was correctly predicted (Figure 8), and 68% of the sample in need of moderate/maximal assistance was correctly predicted (Figure 9).

Next, I analyzed the accuracy when the predicted ADL motor and ADL process decisions were the same (e.g., ADL motor and ALD process ability measures both fell within the same decision zone). When matching decisions existed (n = 30,835, 48% of the sample), 90% of the independent sample was correctly predicted (Figure 7), 22% of the sample in need of minimal assistance was correctly predicted (Figure 8), and 88% of the sample in need of moderate/maximal assistance was correctly predicted (Figure 9).

Accuracy Across Different Diagnostic Categories

For the last analyses, ROC curve analyses were generated across the two different test conditions (i.e., independence vs. in need of assistance and other vs. moderate/maximal assistance) and the area under the resulting curves were compared.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Due to relatively small sample sizes and invariant functional level ratings, the diagnostic groups of well, mild learning disability, and frail were not analyzed. When utilizing ADL motor ability to determine independence versus in need of assistance and other versus moderate/maximal assistance, the area under the curve ratings ranged from 0.68 to 0.85 and 0.66 to 0.85, respectively, indicating poor to excellent potential to accurately categorize clients within the diagnostic groups (Table 18). When utilizing ADL process ability to determine independence versus in need of assistance and other versus moderate/maximal assistance, the area under the curves ranged from 0.77 to 0.92 and 0.75 to 0.84, respectively, indicating fair to excellent potential to accurately categorize clients within the diagnostic groups.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

CHAPTER 5

DISCUSSION

Summary of Findings

Overall, the analyses indicate that both ADL motor and ADL process ability have utility in their use as indicators of level of community dependence. The ADL process ability continues to be more closely associated to global functional level, as indicated by the higher correlation and overall higher area under the ROC curve statistics. Such findings are in line with previous studies, where the ADL process ability scale was found to have more utility than the motor scale in estimating the need for assistance (Hartman et al., 1999; Fisher, 2006a). Although previous studies have documented the relatively high association between ADL motor ability and level of assistance (A. G. Fisher, personal communication, January 13, 2005), researchers have not previously proposed the use of the ADL motor scale in the determination of the need for assistance to live in the community (Hartman et al., 1999; Fisher 2006a). In contrast, the findings of this study suggest that ADL motor ability does have promise in its use as an indicator of level of community dependence, as demonstrated by the relatively high area under the ROC curve statistics.

Although the current ADL motor cutoff measure of 2.00 logits has not been widely used as an indicator of independence in the community, the examination revealed that the current cut is too high, as evidenced by the large false negative rate (60% of the independent sample was incorrectly coded as needing assistance). In contrast, the current

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

ADL process cutoff measure of 1.00 logits was relatively accurate; however, the false positive rate was a bit higher than desired, with 30% of the sample in need of assistance being incorrectly identified as independent.

Prior to beginning the analyses, I determined criteria for selecting the optimal cutoff measure to be the point on the scale that minimized the false positive rate, without resulting in a large decrease in the true positive rate. Based on this criterion, I investigated the possibility of creating new ADL motor and ADL process cutoff measures. Utilizing the output of the ROC curve analysis, new cutoff measures were set at 1.50 and 1.10 logits on the ADL motor and ADL process scales, respectively. Thus, clients with ability measures above the ADL motor and/or ADL process cutoff measures were predicted to be independent and those with ability measures below the ADL motor and/or ADL process cutoff measures were predicted to be in need of assistance to live in the community.

To see if I could improve the accuracy of the proposed cutoff measures, I determined the accuracy of the decisions when the predicted ADL motor and ADL process decision points matched. Examining the use of matched ADL motor and ADL process decisions has not been done prior to this study, and thus the findings have the potential to provide the clinical and research communities with a new means of examining the constellation of ability measures, with relation to estimating the need for assistance to live in the community. Overall, the highest degree of accuracy occurred when the ADL motor and ADL process decisions matched, with false negative and false positive error rates both dropping to 17%. Although the error rates were lowest when the ADL motor and ADL process decision points matched, it is important to note that when

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

discrepancies arise between the motor and process decision points, the better predictor of the need for assistance is the new ADL process ability cutoff measure.

Next, I utilized ROC curve analyses to determine if the ADL motor and ADL process ability measures could be utilized to accurately categorize individuals who need moderate to maximal assistance to live in the community. Based on the fair to excellent area under the ROC curve ratings (Table 17), both the ADL motor and ADL process ability measures have utility in determining the need for moderate to maximal assistance, with ADL process ability once again proving to be the stronger indicator. The addition of the moderate/maximal assistance cutoff measures provides the clinical and research communities with new evidence and new guidelines for estimating the *amount* of assistance needed.

The new cutoff measures demarcating the need for moderate to maximal assistance were set at 1.10 logits and 0.70 logit on the ADL motor and ADL process scales, respectively. Interestingly, the new ADL process cutoff measure falls at the bottom edge of the risk zone previously defined as the area between 1.30 and 0.70 logits on the process scale (Bernspång & Fisher, 1995a; Hartman et al., 1999).

As above, to see if I could improve the accuracy of the predicted decisions, I examined the accuracy of the moderate/maximal assistance cutoff measures when the new ADL motor and ADL process decision points matched. When matched decisions occurred, the false negative rate dropped to 12% and the false positive rate dropped to 23%. Once again, although the matched decision points resulted in the lowest error rates, if the decision points did not match, the new ADL process cutoff measure was the better indicator of the need for moderate/maximal assistance.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

In the next stage of the project, I examined the accuracy of using multiple cutoff measures on the ADL motor and ADL process scales to categorize individuals into one of three different functional level categories (independent, in need of minimal assistance, or in need of moderate to maximal assistance). The ability to correctly categorize individuals who are either independent in the community or in need of moderate to maximal assistance to live in the community is overall quite good. At best, when ADL motor and ADL process decision points matched, 90% of the independent sample and 88% of the sample in need of moderate to maximal assistance were accurately predicted (Figures 7 & 9). Despite the success with the independent and moderate to maximal assistance groups, the ability to accurately categorize those in need of minimal assistance was lower than expected; at best, only 33% of the minimal assistance group was correctly categorized.

Lastly, I globally explored potential differences in the accuracy of the ADL motor and ADL process ability measures across different diagnostic categories. Although the relative accuracy of the ADL motor and ADL process ability measures across the different diagnostic categories varies between the two categorization conditions (i.e., independent vs. in need of assist and other vs. moderate/maximal assist), a few trends in the data were noted. For example, when comparing the ADL motor scale and the ADL process scale across the different diagnostic categories, I surprisingly found that for both categorization conditions, the ADL motor scale was relatively more accurate than the ADL process scale when assessing clients with medical and musckuloskeletal diagnoses. Otherwise, the ADL process scale was consistently the best predictor. Based on the literature I reviewed, this study provides the first line of evidence that the ADL motor

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

scale may have relatively more utility than the ADL process scale in determining the need for assistance and level of community dependence for persons with medical or musculoskeletal disorders.

Another interesting finding was that the estimations were less accurate overall for clients with psychiatric diagnoses. Although such findings are interesting and noteworthy, one must understand that the area under the ROC curve is only a global indicator of an assessment's ability to correctly categorize individuals into one of two groups. More importantly, the area under the ROC curve does not provide conclusive evidence of the assessment's utility and accuracy, and thus one must use caution when considering the results of this analysis. Regardless, based on the differing area ratings obtained, there is evidence to support continued exploration of the accuracy of the proposed cutoff measures within and between different diagnostic categories.

Interpretation of the Results

The most perplexing factor in this study was the relative inability to utilize the AMPS ADL motor and/or ADL process ability measures to categorize individuals who need minimal assistance to live in the community. Reasons for this are likely multifaceted and hard to precisely determine. One explanation lies within the fact that the AMPS was not designed to determine level of dependence in the community, thus, a certain amount of inaccuracy and error is expected. It is known that ADL ability and global functional level are related (medium effect), but the correlation is far from perfect, and, thus, some inaccuracy in predicting level of community dependence is expected.

To explore this concept further, consider a client who can independently and effectively complete PADL and IADL tasks, yet needs assistance with tasks that have

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

social interaction expectations (e.g., public transportation, leisure, productive occupations). Since the AMPS is an assessment ADL motor and ADL process skills during the performance of PADL and IADL tasks, it cannot be used to assess social interaction skills that are often required during the performance of social tasks. Thus, it is conceptually possible that a client with this ADL ability profile, with both AMPS ADL ability measures above the upper cutoff measures incorrectly indicating independence in the community, would actually need some assistance. In the example above, the occupational therapist would need to gather and utilize additional information, beyond the AMPS, to determine the most accurate level of community dependence. Hence, the AMPS ADL ability measures cannot be expected to account for all of the variation in level of community dependence.

Explaining variation in level of community dependence may change as new and diverse demands are placed upon individuals in society, especially if they are not related to ADL motor and/or ADL process ability. For example, individuals in society are being called upon to interact more and more with computer technology and virtual environments. In any given day, I may find myself paying bills on the computer, getting cash from the automated teller machine, purchasing goods with a debit card machine, ordering prescriptions over the phone or on-line, and staying in touch with family and friends via email. With this, I ask myself, does one need to have (or will one need to have) basic computing/technological skills to be independent in the community? A recent study examined the validity and reliability of the Assessment of Computer-Related Skills (ACRS), and although some of the motor and process skills of the AMPS were adapted to fit under the construct of computer skills, it is not known if the ACRS

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

measures a different construct than AMPS ADL motor and ADL process ability (Fischl & Fisher, in press). Thus, the possibility exists that there are other constructs of daily life performance (e.g., social interaction, computer or technology skills) that could explain some of the variation in level of community dependence.

Another area to explore, with regard to explaining the potential reasons for the error in the predictions, is rater scoring error. One scoring error that can occur, amongst occupational therapists working with clients with physical disabilities, is an over emphasis of observing and scoring ADL motor skill deficits and the simultaneous failure to observe and accurately score ADL process skill deficits. Likewise, rater scoring error can also occur amongst occupational therapists who work primarily with clients who have cognitive or psychiatric deficits. Here, the raters may focus their observations and scoring on ADL process skills, while failing to observe and accurately score ADL motor skill deficits. Although physical and cognitive/mental impairments cannot be equated to ADL skill (see Chapter 2), it is possible that such rater error may occur when occupational therapists have not (a) grasped the difference between ADL skill and body functions, (b) adequately developed their skills in simultaneously observing both ADL motor and ADL process skill, and/or (c) separated their preconceived understandings or biases from their actual observations.

The first scoring error noted above could result in inflated ADL process ability measures and decreased utility of using the ADL process scale when predicting level of dependency. This may offer one explanation as to why the ADL process scale is relatively less accurate than the ADL motor scale in the sample of clients with musculoskeletal and medical disorders (Table 18). In contrast, the second error noted

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

above could result in inflated ADL motor ability measures and decreased utility of using the ADL motor scale when predicting level of care.

Another source of error within the current study may be the accuracy of the occupational therapists' ratings of global functional level. First, the reliability and validity of the AMPS global functional level ratings have not been formally evaluated. Therefore, we must proceed with some caution. The possibility exists, for example, that clinicians are more accurate when determining the global functional levels of those who are on the extreme ends of the scale (e.g., clients who are clearly independent and those who require substantial assistance). This may explain the higher rate of accurate predictions for the independent and moderate to maximal assistance groups. On the other hand, when clients require some assistance, and/or when there are inconsistent needs for assistance, the possibility exists that the clinician's judgments are inaccurate and/or inconsistent.

Experience with the AMPS and rater training has supported the possibility that some occupational therapists do not score the global functional level ratings accurately. One common source of error can occur when the occupational therapist does not recognize that a spouse or caregiver's support is what enables a client to live "independently" in the community. In such instances, the occupational therapist incorrectly rates the client as being independent in the community, when in fact he or she would not be able to live at home without support and assistance. Another source of error arises when a client, who lives alone, is assumed to be independent in all aspects of community living. It is important to note that even though a client lives alone, he or she may not be able to independently and/or safely complete all of the tasks necessary for

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

independent living. For example, I have personally worked with many older adults who lived alone, yet were unable to independently and/or safely complete routine household chores (e.g., clean the bath tub, vacuum the carpet, mop the floor), resulting in unsanitary and unsafe living environments. Such clients, although they lived alone, fell below the upper ADL motor and/or ADL process ability cutoff measures on the AMPS Graphic Report.

To continue exploring the potential inaccuracies in predicting global functional level, accurately assessing the needs of clients with psychiatric illness has been shown to be quite difficult (Cheah, Parker, & Roy, 2000; Cheah, Parker, Hadzi-Pavlovic, Gladstone & Eyers, 1998). For example, Chea and colleagues have documented that there is little agreement between (a) the self-reported needs of those with mental illness, and (b) those reported by caregivers and professional staff members. Additionally, they stressed that the concept of need for services and assistance has been inconsistently defined within the psychiatric community, and therefore differing, views of the exact nature of the need for services and assistance have arisen. As a result, the estimates of needed supports are less precise. Consequently, decreased accuracy in predicting level of community dependence for clients with psychiatric illnesses may be related to inaccurate ratings the global functional level provided by the occupational therapist versus decreased sensitivity of the AMPS ability measures.

Finally, 47% of the psychiatric sample was rated as needing minimal assistance by the assessing occupational therapist. Compared to the other diagnostic groups in the sample, the psychiatric diagnostic group had the highest percentage of clients in need of minimal assistance. At this point, one may think that the persons in the sample of

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

psychiatric clients truly need minimal assistance to live in the community, leading to the notion that the AMPS ability measures are less accurate when used to predict the need for minimal assistance. On the other hand, due to the inherent difficulties determining their client's precise needs for assistance, one may also wonder if the occupational therapists in mental health settings had higher rates of error when documenting global functional level. In this instance, the error may be more related to the rating of the global functional level versus the inability of the AMPS to determine level of dependency.

The possibility also exists that the 3-point rating scale of the AMPS global functional level is not sensitive to changes in ability. Since moderate and maximal assistance to live in the community have been combined into the same global functional level rating, the possibility of inaccurate ratings exists. For example, if a therapist believes that a client needs more than minimal assistance, but not maximal assistance to live in the community, there may be an inclination to document the need for minimal assistance rather than moderate/maximum assistance. If there was a fourth global functional level (e.g., the need for moderate assistance), perhaps the clinician's judgments would be more accurate/precise, which could lead to more accurate estimates of level of community dependence based on the AMPS ability measures. Conversely, it may be that the inclusion of a fourth rating level could lead to more error, due to the difficulties in differentiating those who need moderate assistance and those who need maximal assistance. Further research could explore the reliability and validity of the AMPS global functional level rating system.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Clinical Recommendations

From a clinical perspective, the primary goal of embarking on this study was to determine if the AMPS, an ADL assessment, could provide occupational therapists with an indication of how much assistance their clients likely need to live in the community. As discussed above, since ADL assessments have not been specifically designed to assess level of dependence in the community, it is unlikely that any ADL assessment would be a perfect predictor of the amount of assistance needed for community living. Despite this, there is solid evidence that the AMPS, and the resulting ADL motor and ADL process ability measures, can provide the occupational therapist with evidence to support their recommendations of how much assistance their clients need.

Although some of the predictions are relatively accurate when utilizing AMPS ADL motor and/or ADL process ability measures to determine level of community dependence, it is important to stress that such decisions are rarely made based on the results of one assessment. For example, when called upon to make discharge recommendations, occupational therapists often (a) assess performance in more than one occupational area, (b) gather pertinent information with regard to the client's physical and social environments, (c) formally or informally assess social interaction skills, and/or (d) gather information from other healthcare providers to make the most accurate and client-centered decisions.

In support of this notion, researchers found that clinicians arrive at discharge recommendations through the utilization of gathering information from multiple sources (Jette, Grover, & Keck, 2003). In their study, Jette and colleagues found that occupational therapists primarily utilize functional assessments to determine discharge

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

needs. However, to arrive at their final discharge recommendations, therapists also consider the patient's wants and needs, the patient's life context, the severity and prognosis of the client's impairment, and other professional opinions. Additionally, all of this information is filtered through the clinician's level of experience, with experienced therapists being more willing to allow the client to take greater risks. Given this information, within the following recommendations it is important to stress that I would never recommend that a therapist utilize *only* the results of the AMPS to document level of community dependence. Rather, ADL motor and ADL process ability measures should serve to contribute to and support the clinician's clinical decisions and recommendations.

When using the AMPS Graphic Report to support clinical recommendations of level of care needed, the clinician can have the most confidence in their decisions when both the ADL motor and ADL process ability measures indicate independence or when both ability measures indicate the need for moderate to maximal assistance. If matched decisions do not occur, the next best option is to utilize the ADL process scale to determine level of community dependence. As the accuracy of documenting the need for minimal assistance to live in the community is quite low, the zone in between the two ADL cutoff measures should be considered and named a "risk zone" (Appendix D). Therefore, if both of the ADL ability measures fall in the risk zone, or if only the ADL process ability measure falls in the risk zone, there is inconclusive evidence of how much assistance the client likely needs to live in the community. Therefore, if this situation arises, it becomes even more critical for the occupational therapist to draw upon additional clinical information to determine how much assistance the client likely needs.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Not only can the AMPS ability measures be utilized to help determine the level of assistance required, the occupational therapist can also utilize the measures to document the need for continued therapy. Consider the following situation; a client admitted to the hospital obtains initial ADL motor and ADL process ability measures of 0.90 and 0.01 logits respectively. Thus, both ADL motor and ADL process ability measures indicate the likely need for moderate to maximal assistance. Additionally, the occupational therapist noted that prior to admission to the hospital, the client lived independently in a suburban home. Given this information, we can conclude that he is not at his functional baseline of independence, and thus there is likely a need for continued therapy. Therefore, the occupational therapist can use the results of the initial AMPS evaluation to plan and implement occupational therapy interventions. As his rehabilitation continues, we may see that his ADL ability improves to baseline status of independence, in which case we have evidence that he is likely able to return home and that he has benefited from our interventions. If, on the other hand, we begin to see that his ADL abilities plateau at a level that does not indicate that he is likely independent in the community, and that continued intervention does not appear to be affecting his occupational performance, we may determine that it is time to discharge the client with the needed supports. The final AMPS evaluation could thus be utilized to document and support our clinical recommendations of how much assistance he likely needs. In the above example, the occupational therapist can utilize the results of the AMPS evaluations to support clinical decisions and to document (a) the need for occupational therapy intervention, (b) the client's ability to return home to independent living, and/or (c) the need for continued assistance and/or discharge recommendations.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Objective Measurement

As discussed in Chapter 2, for a measure to be considered truly objective, social objectivity must be developed through the work of metrology (Fisher 2000, 2002). To review, metrology is a term used to describe the purposeful development of symbol–meaning coordination. In other words, it is the widespread development of a "common language" regarding the meaning of measures. In order to understand the meaning of a measure, one must first compare the measure to a well-understood concept/construct. Although the aim of the current study was not to define the concepts of ADL motor and ADL process ability, I have provided an expanded means of interpreting the measures and have thus provided a small avenue towards the creation of a common language. If an occupational therapist utilizes the AMPS to support their clinical recommendations, he/she must educate others as to the meanings behind the measures. Through education and the utilization of the AMPS ADL measures, it will become common knowledge within clinical settings of what the measures indicate (e.g., varying levels of assistance, effort, efficiency, and/or safety), and social objectivity will emerge.

To explain further, now that the reader has had the opportunity to read the results (including tables and figures) and discussion within this dissertation, the reader has the tools and knowledge to begin to interpret the meaning of the ADL motor and ADL process ability measures. Prior to reading this paper, if I had stated that my client's ADL motor and ADL process ability measures were 0.80 and 0.50 logits, respectively, the reader would not have an understanding of what these measures mean, as there was not a common understanding of what the measures represent. However, through reading this research paper the reader now has the tools (i.e., Appendix B and D) and knowledge to

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

generate a general idea of the ability of this person. Although this is only one piece of information, we could deduce that the client is likely in need of moderate/maximal assistance to live in the community, that he is experiencing increased effort, decreased efficiency, and possibly decreased safety during the performance of familiar ADL tasks. Within this, depending on the client's baseline status and unique circumstances, either we would start to make plans to continue therapy to improve the client's ADL ability or we would begin to arrange for caregiver assistance upon discharge. If, as the reader of this dissertation, you came to the above conclusions and clinical options when you noted the client's ADL ability measures, then I have been successful in creating symbol–meaning coordination and social objectivity among the readers of this paper. The next step must come from the widespread use and education of the AMPS ADL ability measures in clinical practice so that the rehabilitation community achieves strong symbol–meaning coordination of the AMPS ADL ability measures.

Potential Limitations of the Study and Recommendations for Future Research

Although clinical judgment shows promise in the determination of global functional level, one potential limitation of this study was the lack of evidence supporting the validity/reliability of the AMPS global functional level ratings. Regardless, this was not considered a major limitation of the study; rather the utilization of the AMPS global functional level ratings allowed us the opportunity to utilize a large data set to generate evidence of the use of the AMPS ADL ability measures in the determination of level of community dependence.

One may also question whether the AMPS global functional level ratings were biased by the AMPS evaluation (or visa versa). First, it is important to note that when

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

rating the client's global functional level, the AMPS raters were (a) blind to the person's final AMPS ability measures, and (b) not aware that the functional level ratings would be used to investigate the validity of the AMPS ability measures in predicting level of care. Additionally, there is no indication within the AMPS manual or the professional literature regarding where individuals of varying functional levels fall on the ADL motor and/or ADL process ability scales. As a result, the clinicians were naïve to how the global functional level ratings would affect the determination of specific levels of community independence/dependence along the continua of linear ADL motor and ADL process ability measures. Thus, I have concluded that the potential for biased ratings was actually quite small, and not considered a limitation of the study.

Future studies could seek to verify the results of the current study. One future research project would be to cross validate the proposed ADL motor and ADL process cutoff measures by diagnostic category, as discussed above. Future research could also seek to validate the proposed cutoff measures through comparing qualitative reports of the clinician's discharge recommendations of the amount of assistance needed to the AMPS ability measures, and later verifying all information against how much assistance the client actually needed upon return home. Lastly, as standardized assessments of global functional level do not appear to exist, future research could examine the accuracy of the AMPS global functional levels ratings via utilizing qualitative methods of inquiry.

Conclusion

In summary, I have concluded that there is merit in utilizing ADL motor and ADL process ability measures to *contribute* to the documentation and determination of the need for assistance to live in the community. In the analysis, the cutoff measures

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

demarcating independence were set at 1.5 logits on the ADL motor scale and 1.1 logits on the ADL process scale; individuals with ability measures above these cuts are likely independent, whereas, those with ability measures below these cuts likely need assistance to live in the community. The cutoff measures demarcating moderate/maximum cuts were set at 1.0 and 0.70 logits on the motor and process scales; individuals with ability measures below these cuts likely need moderate to maximal assistance to live in the community. Overall, the best indicator of independence occurred when both ability measures were above the independent cuts. Likewise, our best indicator of the need for moderate to maximal assistance occurred when ADL motor and ADL process ability measures were below the lower cutoff measures. When discrepancies occur between the ADL motor and ADL process decision points, the most accurate decisions arise from the use of the ADL process cutoff measures.

The zone between the upper and lower cutoff measures was termed the "risk zone," as there is inconclusive evidence of *how* much assistance individuals need if they have ability measures within this zone. If both ADL motor and ADL process ability measures fall with the risk zone, or if the ADL process ability measure is in the risk zone, the occupational therapist must gather additional information to make accurate decisions of level of community dependence.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

REFERENCES

Agran, M., Marchand-Martella, N. E., & Martella, R. C. (Eds.). *Promoting health and safety: Skills for independent living.* Baltimore, MD: Paul H. Brooks.

American Occupational Therapy Association (2002). Occupational therapy practice framework: Domain and process. *American Journal of Occupational Therapy, 56,* 609-639.

Au, R. W. C., Tam, P. W. C., Tam, G. W. C., & Ungvari, G. S. (2005). Cross-cultural validation of the St. Louis Inventory of Community Living Skills for Chinese patients with schizophrenia in Hong Kong. *Psychiatric Rehabilitation Journal, 29,* 34-40.

Åsberg, K.H. & Nydevik, I. (1991). Early prognosis of stroke outcome by means of Katz Index of Activities of Daily Living. *Scandinavian Journal of Rehabilitation Medicine, 23,* 187-191.

Bachelder, J. (1985). Independent living programs: Bridges from hospital to community. *Occupational Therapy in Health Care, 2,* 99-107.

Bennett, S. & Bennett, J. W. (2000). The process of evidence-based practice in occupational therapy: Informing clinical decisions. *Australian Occupational Therapy Journal, 47,* 171-180.

Berk R. A. (ed.). (1984). *A guide to criterion-referenced test construction.* Baltimore, MD: John Hopkins University Press.

Bernspång, B., & Fisher, A. G. (1995a). Differences between persons with right or left CVA on the Assessment of Motor and Process Skills. *Archives of Physical Medicine and Rehabilitation, 76,* 1144-1151.

Bernspång, B., & Fisher, A. G. (1995b). Validation of the Assessment of Motor and Process Skills for use in Sweden. *Scandinavian Journal of Occupational Therapy, 2,* 3-9.

Biagini, R. E., Krieg, E. F., Pinkerton, L. E., & Hamilton, R. G. (2001). Receiver operating characteristic analyses of food and drug administration-cleared serological assays for natural rubber latex-specific immunoglobulin E antibody. *Clinical and Diagnostic Laboratory Immunology, 8,* 1145-1149.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Bishop, C. E., Gilden, D., & Blom, J. (2002). Medical spending for injured elders: Are there opportunities for savings? *Health Affairs, 21,* 215-223.

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

Brorsson, B. & Åsberg, K.H. (1984). Katz Index of Independence in ADL: Reliability and validity in short-term care. *Scandinavian Journal of Rehabilitation Medicine, 16,* 125-132.

Cheah, Y.-C., Parker, G., Hadzi-Pavlovic, D., Gladstone, G., & Eyers, K. (1998). Development of a measure profiling problems and needs of psychiatric patients in the community. *Social Psychiatry and Psychiatric Epidemiology, 33,* 337-344.

Cheah, Y.-C., Parker, G., & Roy, K. (2000). Evaluation and validation of a measure profiling needs and problems of psychiatry patients in the community: A Malaysian study. *Social Psychiatry and Psychiatric Epidemiology, 35,* 170-176.

Choo, B., Umraw, N., Gomez, M., Cartotto, R. & Fish, J. S. (2006). The utility of the Functional Independence Measure (FIM) in discharge planning for burn patients. *Burns, 32,* 20-23.

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

Cooke, K. Z., Fisher, A. G., Mayberry, W., & Oakley, F. (2000). Differences in activities of daily living process skills of persons with and without Alzheimer's disease. *Occupational Therapy Journal of Research, 20,* 87-104.

Cotter, E. M., Burgio, L. D., Stevens, A. B., Roth, D. L., & Gitlin, L. N. (2002). Correspondence of the Functional Independence Measure (FIM) self-care subscale with real-time observations of dementia patients' ADL performance in the home. *Clinical Rehabilitation, 16*, 3-45

Cry, M., Toupin, J. Lessage, A. D., Valiquette, C. M. (1994). Assessment of independent living skills for psychotic patients: Further validity and reliability. *Journal of Nervous and Mental Disease, 182*(2), 91-97.

Darragh, A. R., Sample, P. L., & Fisher, A. G. (1998). Environment effect on functional task performance in adults with acquired brain injuries: Use of the Assessment of Motor and Process Skills. *Archives of Physical Medicine and Rehabilitation, 79,* 418-423.

DeBettignies, B. H. & Mahurin, R. K. (1989). Assessment of independent living skills in geriatric populations. *Clinics in Geriatric Medicine, 5,* 461-475.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

DeLoach, C. P., Wilkins, R. D., & Walker, G.W. (1983). *Independent living: Philosophy, process, and services.* Balitimore, MD: University Park Press.

Desrosiers, J., Bravo, G., Hébert, R., & Dubuc, N. (1995). Reliability of the revised Functional Autonomy Measurment System (SMAF) for epidemiological research. *Age and Ageing, 24,* 402-406.

Desrosiers, J., Rochette, A., Noreau, L., Bravo, G., Hébert, R., & Boutin, C. (2003). Comparison of two functional independence scales with a participation measure in post-stroke rehabilitation. *Archives of Gerontology and Geriatrics, 37,* 157-172.

Deutsch, A., Braun, S., & Granger, C. (1996). The Functional Independence Measure (FIM^{SM}) and the Functional Independence Measure for children (WeeFIM® Instrument): Ten years of development. *Critical Reviews in Physical Rehabilitation Medicine, 8*, 267-281.

Dever, R. B. & Knapczyk, D. R. (1988). *Community living skills: A taxonomy.* Washington, DC: American Association on Mental Retardation.

Dickerson, F. B. (1997). Assessing clinical outcomes: The community functioning of persons with serious mental illness. *Psychiatric Services, 48*, 897-902.

Dickerson, A. E., & Fisher, A. G. (1993). Age differences in functional performance. *American Journal of Occupational Therapy, 47*, 686-692.

Dickerson, A. E., & Fisher, A. G. (1995). Culture-relevant functional performance assessment of the Hispanic elderly. *Occupational Therapy Journal of Research, 15*, 50-68.

Dijkstra, A., Tiesinga, L. J., Plantinga, L., Veltman, G., & Dassen, T. W. (2005). Diagnostic accuracy of the care dependency scale. *Journal of Advanced Nursing, 50,* 410-416.

Disler, P.B., Roy, C.W., & Smith, B.P. (1993). Predicting hours of care needed. *Archives of Physical Medicine and Rehabilitation, 74*, 139-143.

Doble, S. E., Fisk, J. D., Fisher, A. G., Ritvo, P. G., & Murray, T. J. (1994). Functional competence of community-dwelling persons with multiple sclerosis using the Assessment of Motor and Process Skills. *Archives of Physical Medicine and Rehabilitation, 75*, 843-851.

Doble, S. E., Fisk, J. D., MacPherson, K. M., Fisher, A. G., & Rockwood, K. (1997). Measuring functional competence in older persons with Alzheimer's disease. *International Psychogeriatrics, 9,* 25-38.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Donkervoort, M., Dekker, J. Stehmann-Saris, F. C., & Deelman, B. G. (2001). Efficacy of strategy training in left hemisphere stroke patients with apraxia: A randomised clinical trial. *Neuropsychological Rehabilitation, 11,* 549-566.

Dromerick, A. W., Edwards, D. F., & Diringer, M. N. (2003). Sensitivity to changes in disability after stroke: A comparison of four scales useful in clinical trials. *Journal or Rehabilitation Research and Development, 40,* 1 – 8.

Dubuc, N., Haley, S. M., Kooyoomjian, J. T., & Jette, A. M. (2004). Assessing disability in older adults: The effects of asking questions with and without health attribution. *Journal of Rehabilitation Medicine, 36,* 226-231.

Dubuc, N., Hébert, R., Desrosiers, J., Buteau, M., & Trottier, L. (2005). Disability-based classification system for older people in integrated long-term care services: The Iso-SMAF profiles. *Archives of Gerontology and Geriatrics, 42,* 191-206.

Duran, L., & Fisher, A. G. (1996). Male and female performance on the Assessment of Motor and Process Skills. *Archives of Physical Medicine and Rehabilitation, 77,* 1019-1024.

Egan, J. P. (1975). *Signal detection theory and ROC analysis.* New York, NY: Academic.

England, W. L. (1988). An exponential model used for optimal threshold selection on ROC curves. *Medical Decision Making, 8,* 120-131.

Evenson, R. C. & Boyd, M. A. (1993). The St. Louis Inventory of Community Living Skills. *Psychological Rehabilitation Journal, 17*(2), 93-99.

Fischl, C. & Fisher, A. G. (in press). Development and Rasch analysis of the Assessment of Computer-Related Skills. *Scandinavian Journal of Occupational Therapy.*

Fisher, A. G. (1993). The assessment of IADL motor skills: An application of many-faceted Rasch analysis. *American Journal of Occupational Therapy, 47,* 319-329.

Fisher, A. G. (1994). Development of a functional assessment that adjusts ability measures for task simplicity and rater leniency. In M. Wilson (Ed.), *Objective measurement: Theory into practice* (Vol 2, 145-175). Norwood, NJ: Ablex.

Fisher, A. G. (1997). Multifaceted measurement of daily life task performance: Conceptualizing a test of instrumental ADL and validating the addition of personal ADL tasks. *Physical Medicine and Rehabilitation: State of the Art Reviews, 11,* 289-303.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Fisher, A. G. (1998). Uniting practice and theory in an occupational framework, 1998 Eleanor Clarke Slagle Lecture. *American Journal of Occupational Therapy, 52,* 509-521.

Fisher, A. G. (2006a). *Assessment of Motor and Process Skills. Vol. 1: Development, standardization, and administration manual* (6th ed.) Fort Collins, CO: Three Star Press.

Fisher, A. G. (2006b). *Assessment of Motor and Process Skills. Vol. 2: User manual* (6th ed.) Fort Collins, CO: Three Star Press.

Fisher, A. G. (2006c). Overview of performance skills and client factors. In H. M. Pendleton & W. Schultz-Krohn (Eds.), *Pedretti's occupational therapy: Practice skills for physical dysfunction* (6th ed., pp. 372-402). St. Louis MO: Mosby Elsevier.

Fisher, A. G., Liu, Y., Velozo, C. A., & Pan, A. W. (1992). Cross-cultural assessment of process skills. *American Journal of Occupational Therapy, 46,* 876-885.

Fisher, W. P. (2000). Objectivity in psychosocial measurement: What, why, how. *Journal of Outcome Measurement, 4,* 527-563.

Fisher, W. P. (2002). *Deconstructing science's mathematical metaphysics: Measurement and the creation of capital.* Unpublished manuscript.

Fitz, D. & Evenson, R. C. (1995). Brief report: A validity study of the St. Louis Inventory of Community Living Skills. *Community Mental Health Journal, 31,* 369-377.

Fortinsky, R. H., Iannuzzi-Sucich, M., Baker, D. I., Gottschalk, M., King, M. B., Brown, C. J., & Tinetti, M. E. (2004). Fall-risk assessment and management in clinical practice: Views from healthcare providers. *Journal of the American Geriatrics Society, 52,* 1522-1526.

Foster, C. B., Gorga, D., Padial, C., Feretti, A. M., Berenson, D., Kline, R., BeLue, R., & Charlson, M. E. (2004). The development and validation of a screening instrument to identify hospitalized medical patients in need of early functional rehabilitation assessment. *Quality of Life Research, 14,* 1099-1108.

Forette, F, Panisset, M., & Boller, F. (1992). Clinical trials in cognitive impairment in the elderly. *Aging: Clinical Experimental Research, 4,* 239-250.

Gair, G. & Hartery, T. (2001). Medical dominance in multidisciplinary teamwork: a case study of discharge decision-making in a geriatric assessment unit. *Journal of Nursing Management, 9,* 3-11.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Girard, C., Fisher, A. G., Short, M. A., & Duran, L. (1999). Occupational performance differences between psychiatric groups. *Scandinavian Journal of Occupational Therapy, 6,* 119-126.

Globe Fearon Educational Publisher. (1997). *Skills for independent living.* Upper Saddle River, NJ: Author.

Goldman, S. L., & Fisher, A. G. (1997). Cross-cultural validation of the Assessment of Motor and Process Skills (AMPS). *British Journal of Occupational Therapy, 60,* 77-85.

Goto, S., Fisher, A. G., & Mayberry, W. L. (1996). AMPS applied cross-culturally to the Japanese. *American Journal of Occupational Therapy, 50,* 798-806.

Graff, M. J. L., Vernooij-Dassen, M. J. F. J., Hoefnagesl, W. H. L., Decker, J. & de Witte, L. P. (2003). Occupational therapy at home for older individuals with mild to moderate cognitive impairments and their primary caregivers: A pilot study. *Occupational Therapy Journal of Research, 23,* 155-164.

Granger, C.V., Cotter, A.C., Hamilton, B.B., & Fiedler, R.C. (1993). Functional assessment scales: A study of persons after stroke. *Archives of Physical Medicine and Rehabilitation, 74,* 133-138.

Granger, C.V., Cotter, A.C., Hamilton, B.B., Fiedler, R.C., & Hens, M. M. (1990). Functional assessment scales: A study of persons with multiple sclerosis. *Archives of Physical Medicine and Rehabilitation, 71,* 870-875.

Granger, C.V., Divan, N., & Fiedler, R.C. (1995). Functional assessment scales: A study of persons after traumatic brain injury. *American Journal of Physical Medicine and Rehabilitation, 74,* 107-113.

Green, J. & Britten, N. (1998). Qualitative research and evidence based medicine. *British Medical Journal, 316,* 1230-1232.

Greiner, M., Pfeiffer, D., Smith, R. D. (2000). Principles and practical application of the receiver-operating characteristic analysis for diagnostic tests. *Preventative Veterinary Medicine, 45,* 23-41.

Gross, D. P. & Battié, M. C. (2006). Does functional capacity evaluation predict recovery in workers' compensation claimants with upper extremity disorders? *Occupational and Environmental Medicine, 63*, 401-410.

Grove, W. M. & Meehl, P. E. (1996). Comparative efficiency of informal (subjective, impressionistic) and formal (mechanical, algorithmic) prediction procedures: The clinical-statistical controversy. *Psychology, Public Policy, and Law, 2,* 293-323.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Haentjens, P., Autier, P., Barette, M.,& Boonen, S. (2001). The economic cost of hip fractures among elderly women: A one-year, prospective, observational cohort study with matched-pair analysis. *Journal of Bone and Joint Surgery, 83-A,* 493-500.

Hall, K. M., Bushnik, T., Lakisic-Kazazic, B., Wright, J., & Cantagallo, A. (2001). Assessing traumatic brain injury outcome measures for long-term follow-up of community-based individuals. *Archives of Physical Medicine and Rehabilitation, 82,* 367-374.

Hanley, J. A. (1989). Receiver operating characteristic (ROC) methodology: The state of the art. *Critical Review of Diagnostic Imaging, 29,* 307-335.

Hariz, G. M., Bergenheim, T., Hariz, M. I. & Lindberg, M. (1998). Assessment of ability/disability in patients treated with chronic thalamic stimulation for tremor. *Movement Disorder, 13,* 78-83.

Hartman, M. L., Fisher, A. G., & Duran, L. (1999). Assessment of functional ability of people with Alzheimer's disease. *Scandinavian Journal of Occupational Therapy, 6,* 111-118.

Hayase, D., Mosenteen, D. A., Thimmaiah, D., Zemke, S., Atler, K., & Fisher, A. G. (2004). Age-related changes in activities of daily living (ADL) ability. *Australian Occupational Therapy Journal, 51,* 192-198.

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

Hébert, R., Carrier, R., & Bilodeau, A. (1988). The Functional Autonomy Measurement System (SMAF): Description and validation of an instrument for the measurement of handicaps. *Age and Ageing, 17,* 293-302.

Hébert, R., Speigalhalter, D. J., & Bayne, C. (1997). Setting the minimal metrically detectable change on disability rating scales. *Archives of Physical Medicine and Rehabilitation, 78,* 1305-1308.

Hébert, R., Guilbault, J., Desrosiers, J., & Dubuc, N. (2001). The functional autonomy measurement system (SMAF): A clinical-based instrument for measuring disabilities and handicaps in older people. *Journal of Canadian Geriatric Society, 4,* 141-147.

Iyer, S. N, Rothmann, T. L., Vogler, J. E., & Spaulding, W.D. (2005). Evaluating outcomes of rehabilitation for severe mental illness. *Rehabilitation Psychology, 50,* 43-50.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Jette, D. U., Grover, L., & Keck, C. P. (2003). A qualitative study of clinical making in recommending discharge placement from the acute care setting. *Physical Therapy, 83,* 224-236.

Katz, S., Downs, T.D., Cash, H.R., & Grotz, R.C. (1970). Progress in the development of the Index of ADL. *Gerontologist, 10,* 20-30.

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

Kinnman, J., Andersson, U., Wetterquist, L, Kinnman, Y., & Andersson, U. (2000). Cooling suit for multiple sclerosis: Functional improvement in daily living? *Scandinavian Journal of Rehabilitation Medicine, 32,* 20-24.

Klein, R. M. & Bell, B. (1982). Self-care skills: Behavioural measurement with the Klein-Bell ADL Scale. *Archives of Physical Medicine and Rehabilitation, 63,* 335-338.

Kottorp, A., Bernspång, B., Fisher, A. G. (2003). Validity of a performance assessment of activities of daily living (ADL) for persons with developmental disabilities. *Journal of Intellectual Disability Research, 47,* 597-605.

Kottorp, A., Bernspång, B., Fisher, A. G., & Bryze, K. (1995). IADL ability measured with the AMPS: Relation to two classification systems of mental retardation. *Scandinavian Journal of Occupational Therapy, 2,* 121-128.

Law, M. & Usher, P. (1988). Validation of the Klein-Bell Activities of Daily Living Scale for children. *Canadian Journal of Occupational Therapy, 55,* 63-68.

Linacre, J. M. (1993). *Many-facet Rasch measurement* (2nd ed.). Chicago: MESA.

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

Linacre, J. M. (2006). Facets Rasch measurement computer program. Chicago: Winsteps.com

Linacre, J. M., Heinemann, A. W., Wright, B. D., Granger, C. V., & Hamilton, B. B. (1994). The structure and stability of the Functional Independence Measure. *Archives of Physical Medicine and Rehabilitation, 75,* 127-132.

Lindén, A., Boschian, K., Eker, C., Schalén, W., Nordström, C. -H. (2005). Assessment of Motor and Process Skills reflects brain-injured patients' ability to resume

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

independent living better than neuropsychological tests. *Acta Neurologica Scandinavica, 111,* 48-53.

Lungren-Nilsson, Å., Grimby, G., Ring, H., Tesio, L., Lawton, G., Slade, A., Penta, M., Tripolski, M., Beiring-Sørensen, F., Carter, J., Marincek, C., Phillips, S., Simone, A., & Tennant, A. (2005). Cross-cultural validity of Functional Independence Measure items in stroke: A study using Rasch analysis. *Journal of Rehabilitation Medicine, 37,* 23-31.

Magalhães, L., Fisher, A. G., Bernspång, B., & Linacre, J M. (1996). Cross-cultural assessment of functional ability. *Occupational Therapy Journal of Research, 16,* 45-63.

McNulty, M. C., & Fisher, A. G. (2001). Validity of using the Assessment of Motor and Process Skills to estimate overall home safety in persons with psychiatric conditions. *American Journal of Occupational Therapy, 55,* 649-655.

Meinow, B., Kåreholt, I., & Lagergren, M. (2005). According to need? Predicting the amount of municipal home help allocated to elderly recipients in urban area of Sweden. *Health and Social Care in the Community, 7,* 366-377.

Mercier, L., Audet, T., Hébert, R., Rochette, & Dubois, M. (2001). Impact of motor, cognitive, and perceptual disorders on ability to perform activities of daily living after stroke. *Stroke, 32*, 2602-2608.

Merritt, B. K., & Fisher, A. G. (2003). Gender differences in the performance of activities of daily living. *Archives of Physical Medicine and Rehabilitation, 84,* 1872-1877.

Michell, J. (1999). *Measurement in psychology: A critical history of a methodological concept.* Cambridge, England: Cambridge University Press.

Michell, J. (2000). Normal science, pathological science and psychometrics. *Theory and Psychology, 10*, 639-667.

Neistadt, M.E. & Seymour, S.G. (1994). Treatment activity preferences of occupational therapists in adult physical dysfunction settings. *American Journal of Occupational Therapy, 49,* 437-443.

Nygård, L., Bernspång, B., Fisher, A. G., & Winblad, B. (1994). Comparing motor and process ability of persons with suspected dementia in home and clinic settings. *American Journal of Occupational Therapy, 48,* 689-696.

Oakley, F., Duran, L., Fisher, A. G., & Merritt, B. (2003). Differences in motor skills in persons with and without Alzheimer's disease. *Australian Occupational Therapy Journal. 50,* 72-78.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Oakley, F., Khin, N., Parks, R., Bauer, L., & Sunderland, T. (2002). Improvement in activities of daily living in elderly following treatment for post-bereavement depression. *Acta Psychiatrica Scandinavica, 105,* 231-234.

Oakley, F., & Sunderland, T. (1997). The Assessment of Motor and Process Skills as a measure of IADL functioning in pharmacologic studies of people with Alzheimer's disease: A pilot study. *International Psychogeriatrics, 9,* 197-206.

Oczkowski, W.J. & Barreca, S. (1993). The Functional Independence Measure: Its use to identify rehabilitation needs in stroke survivors. *Archives of Physical Medicine and Rehabilitation, 74,* 1291-1294.

Pan, A. W., & Fisher, A. G. (1994). The assessment of motor and process skills of persons with psychiatric disorders. *American Journal of Occupational Therapy, 48,* 775-780.

Papaioannou, A., Parkinson, W., Cook, R., Ferko, N., Coker, E., & Adachi, J. D. (2004, January). Prediction of falls using a risk assessment tool in the acute care setting. *BMC Medicine, 2.* Retrieved July 22, 2006, from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=333435

Park, S., Fisher, A. G., & Velozo, C. A. (1994). Using the Assessment of Motor and Process Skills to compare occupational performance between clinic and home settings. *American Journal of Occupational Therapy, 48,* 697-709.

Pepe, M. S. (2003). *The Statistical evaluation of medical tests for classification and prediction.* New York: Oxford University Press.

Perneczky, R., Pohl, C., Sorg, C., Hartman, J., Komossa, K., Alexopoulos, P., et al. (2006). Complex activities of daily living in mild cognitive impairment: Conceptual and diagnostic issues. *Age and Ageing, 35,* 240-245.

Pollak, N., Rheault, W., & Stoecker, J.L. (1996). Reliability and validity of the FIM for persons aged 80 years and above from a multilevel continuing care retirement community. *Archives of Physical Medicine and Rehabilitation, 77,* 1056-1057.

Poole, J. L., Atanasoff, G., Pelsor, J.C., & Sibbett, W. L. (2006). Comparison of a self-report and performance-based test of disability in people with systemic lupus erythematosus. *Disability and Rehabilitation, 28,* 653-658.

Rexroth, P., Fisher, A.G., Merritt, B.K., & Gliner, J. (2005). Ability differences in persons with unilateral hemispheric stroke. *Canadian Journal of Occupational Therapy, 72,* 212-221.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Rice, M. S., Leonard, C., & Carter, M. (1998). Grip strengths and required forces in accessing everyday containers in a normal population. *American Journal of Occupational Therapy, 52,* 621-628.

Robinson, S. E., & Fisher, A. G. (1996). A study to examine the relationship of the Assessment of Motor and Process Skills (AMPS) to other tests of cognition and function. *British Journal of Occupational Therapy, 59,* 260-263.

Robinson, S. E., & Fisher, A. G. (1999). Functional and cognitive differences between cognitively-well people and people with dementia. *British Journal of Occupational Therapy, 62,* 466-471.

Rogers, J. C., Holm, M. G., Beach, S., Schulz, R., Cipriani, J., Fox, A., & Starz, T. W. (2003). Concordance of four methods of disability assessment using performance in the home as the criterion method. *Arthritis and Rheumatism (Arthritis Care & Research), 49,* 640-647.

Rubenstein, L. Z., Schairer, C., Wieland, G., D., & Kane, R. (1984). Systematic biases in functional status assessment of elderly adults: Effects of different data sources. *Journal of Gerontology, 39,* 686-691.

Sackett, D. L., Rosenberg, W. M., Gray, J. A., Haynes, R. B., & Richardson, W. S. (1996). Evidence based medicine: What it is and what it isn't. *British Medical Journal, 312,* 71-72.

Santos-Eggimann, B., Zobel, F., & Bérod A. C. (1998). Functional status of elderly home care users: Do subjects, informal and professional caregivers agree? *Journal of Clinical Epidemiology, 52*, 181-186.

Searight, H. R. & Goldberg, M. A. (1991). The Community Competence Scale as a measure of functional daily living skills. *Journal of Mental Health Administration, 18,* 128-134.

Sellers, S. W., Fisher, A. G., & Duran, L. (2001). Validity of the Assessment of Motor and Process Skills with students who are visually impaired. *Journal of Visual Impairment and Blindness, 95*, 164-167.

Shillam, L. L., Beeman, C., & Loshim, P. M. (1983). Effect on occupationl therapy intervention on bathing independence of disabled persons. *American Journal of Occupational Therapy, 37,* 744 –748.

Smith, E. V. & Smith R. M. (2004). *Introduction to Rasch measurement.* Maple Grove, MN: JAM Press.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Stauffer, L. M., Fisher, A. G., & Duran, L. (2000). ADL performance of black and white Americans on the Assessment of Motor and Process Skills. *American Journal of Occupational Therapy 54,* 607-613.

Stinemann, M.G., Shea, J.A., Jette, A., Tassoni, C.J., Ottenbacher, K.J., Fiedler, R., & Granger, C.V. (1996). The Functional Independence Measure: Tests of scaling assumptions, structure, and reliability across 20 diverse impairment categories. *Archives of Physical Medicine and Rehabilitation, 77,* 1101-1108.

Studenski, S. A., Wallace, D., Duncan, P. W., Rymer, M., & Lai, S. M. (2001). Predicting stroke recovery: Three- and six-month rates of patient-centered functional outcomes on the Orpington Prognostic Scale. *Journal of the American Geriatric Society, 49,* 308-312.

Swets, J. A., Dawes, R. M., & Monahan, J. (2000). Better decisions through science. *Scientific America,* 283, 82-87.

Switzky, H. N., Rotatori, A. F., & Cohen, H. (1978). Community Living Skills Assessment Inventory: An instrument to facilitate deinstitutionalization of the severely developmentally disabled. *Psychological Reports, 43,* 1335-1342.

Tape, T. G. (n.d). Interpreting diagnostic tests. Retrieved January 16, 2007 from http://gim.unmc.edu/dxtests/Default.htm

Tennant, A., Penta, M., Tesio, L., Grimby, G., Thonnard, J., Slade, A. et al. (2004). Assessing and adjusting for cross-cultural validity of impairment and activity limitation scales through differential item functioning within the framework of the Rasch model. *Medical Care, 42*(Suppl. January), 37-48.

Tham, K., Ginsburg, E., Fisher, A. G., & Tegnér, R. (2001). Training to improve awareness of disabilities in clients with unilateral neglect. *American Journal of Occupational Therapy, 55,* 46-54.

Three Star Press (2005). *Assessment of Motor and Process Skills computer-scoring program.* Fort Collins, CO: Author.

Tousignant, M., Hébert, R., Dubuc, N., Simoneau, & Dieleman, L. (2003). Application of a case-mix classification based on the functional autonomy of the residents for funding long-term care facilities. *Age and Aging, 32,* 60-66.

Townsend, E. & Ryan, B. (1991). Assessing independence in the community. *Canadian Journal of Public Health, 82,* 52-57.

Uniform Data System for Medical Rehabilitation (USDMR), 1997. *Guidelines for the Uniform Data Set for Medical Rehabilitation (including the FIM™ instrument), version 5.1.* Buffalo, NY: State University of New York.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Uyttenboogaart, M., Stewart, R. E., Vroomen, P., De Keyser, & Luijckx, G. (2006). Optimizing cutoff scores for the Barthel Index and the Modified Rankin Scale for defining outcome in stroke trials. *Stroke, 36,* 1984-1987.

Van Bennekom, C. A. M., Jelles, F., Lankhorst, G. J., & Bouter, L. M. (1996). Responsiveness of the Rehabilitation Activities and the Barthel Index. *Journal of Clinical Epidemiology, 49,* 39-44.

Wallace, C. J., Boone, S. E., Donahue, P. C., & Foy, D. W. (1985). The chronically mentally disabled: Independent living skills training. In D. H. Barlow (ed.), *Clinical handbook of psychological disorders.* New York, NY: Guilford.

Weir, E. & Culmer, L. (2004). Fall prevention in the elderly population. *Canadian Medical Association Journal, 17,* 724-724.

West, W.L. (1984). A reaffirmed philosophy and practice of occupational therapy for the 1980s. *American Journal of Occupational Therapy, 38,* 15-23.

Whelihan, W. M., Emerson, L. L., Kleban, M. H., & Granick, S. (1984). Mental status and memory assessment as predictors of dementia. *Journal of Gerontology, 39,* 572-576.

Williams, J. H., Drinka, T. J. K., Greenberg, J. R., Farrel-Holtan, J., Euhardy, R., & Schram, M. (1991). Development and testing of the Assessment of Living Skills and Resources (ALSAR) in elderly community-dwelling veterans. *Gerontologist, 31,* 84-91.

Winograd, C. H. (1984). Mental status test and the capacity for self-care. *Journal of the American Geriatric Society, 32,* 49-55.

Webber, L. S., Jenkinson, L. S., & McGillivray, J. A. (2002). Adaptive behaviour in Australia: What items are essential for assessing independent living? *Australian Psychologist, 37,* 63-67.

World Health Organization [WHO] (2001). *International classification of functioning, disability and health: ICF.* Geneva: Author.

Wright, B. D. & Stone, M. H. (1979). *Best test design.* Chicago: MESA press.

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

Wright, B.D. & Stone, M. (1988). Identification of item bias using Rasch measurement. *Research Memorandum 55.* Chicago: MESA; 1988.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Zhou, X., Obuchowski, N. A., McClish, D. A. (2002). *Statistical methods in diagnostic medicine.* New York: John Wiley & Sons.

Zweig, M. H. & Campbell, G. (1993). Receiver-Operating characteristic (ROC) plots: A fundamental evaluation tool in clinical medicine. *Clinical Chemistry, 39,* 561-577.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table 1

The Assessment of Motor and Process Skills (AMPS) Global Functional Levels (Fisher, 2006a, 2006b)

Rating	Definition
Independent	• Able to manage all daily life tasks in the home and the community safely and without assistance • The person has a potential for future suicide or other health safety risk, but currently is able to manage all daily life tasks safely and without assistance • The person is a well child 16 years of age and older
Minimal assistance	• Unable to manage *some* daily life tasks in the home and community safely and without assistance or supervision, such as money management, shopping, community transportation, and heavy home maintenance tasks. • The person requires supervision because of acute suicide or other safety risk, but otherwise is able to manage all daily life tasks without assistance • The person is a well child between 10 and 15 years of age
Moderate to maximal assistance	• Unable to manage all or most daily life tasks in the home and the community safely and without assistance. Needs assistance with daily life tasks such as simple meal preparation, dressing, bathing, or showering • The person is a well child 9 years of age or less

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table 2

Tasks Personal Maintenance and Development Domain (Dever & Knapczyk, 1988)

Personal cleanliness

- Bathe
- Shampoo hair
- Brush/floss teeth
- Maintain clothing cleanliness
- Care for menses (females)
- Clean nails
- Clean nose
- Eliminate waste
- Care for skin
- Other

Grooming

- Trimmed & neat hair
- Comb/brush hair
- Trim nails
- Use deodorant
- Use makeup (females)
- Other

Equipment breakdown or depletion of materials

- Repair/replace broken equipment
- Replace depleted materials
- Perform alternate activities

Dress appropriately

- Neat clothing
- Maintain shoes
- Observe local clothing styles
- Coordinate clothing
- Appropriate clothing for activity
- Appropriate clothing for location (indoor vs. outdoor)
- Appropriate clothing for weather
- Repair/discard worn clothing
- Discard out of style clothing
- Other

Sleep patterns

- Identify required sleep patterns
- Follow required sleep patterns

Medications

- Obtain prescriptions
- Follow prescriptions

Sexual interactions

- Identify sexual needs
- Appropriately satisfy sexual needs

Nutrition

- Eat balanced meals
- Maintain appropriate body weight

Exercise

- Exercise cardiovascular system
- Exercise skeletal muscle system

Illness treatment

- Identify injuries requiring first aid
- Treat minor injuries
- Maintain first aid supplies
- Identify illness requiring home treatment
- Other

Substance control

- Control use of dangerous substances
- Seek assistance if necessary

(continued)

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table 2 (continued)

Family interaction	Interactions with friends	Changes to daily schedule
• Perform required interactions • Refrain from inappropriate interactions • Observe demeanor requirements • Observe conversational constraints • Appropriate responses to emotions of others • Appropriate emotional responses • Display appropriate body language	• Observe required interactions • Refrain from interacting inappropriately • Observe demeanor requirements • Observe conversational constraints • Appropriate responses to emotions of others • Appropriate emotional responses • Appropriate body language	• Cope with sleep pattern disruptions • Cope with the result of sleep pattern disruptions • Perform alternate activities during changes in routine
Obtain necessary assistance with maintaining relationships	Make friends	Obtain medical advice/treatment
• Identify problems in personal relationships • Obtain assistance with personal problems	• Identify potential friends • Appropriate overtures • Make friends	• Identify injury/illness requiring medical intervention • Follow emergency procedures • Follow routine procedures (e.g., periodic examinations)
	Inappropriate family/friend conduct	
	• Identify inappropriate conduct of others • Respond to inappropriate conduct	

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table 3

Tasks in the Domain of Homemaking and Community Life (Dever & Knapczyk, 1988)

Find living quarters

- Locate potential quarters
- Assess desirability of potential living quarters
- Assess ability to pay costs

Buy/rent living quarters

- Select living quarters
- Agree on price
- Sign agreements
- Make deposit
- Budget costs

Set up living quarters

- Obtain furniture
- Obtain food, tools, materials
- Move
- Set up furniture
- Store food, tools, materials

Maintain interior of home

- Paint surfaces
- Repair/replace household equipment
- Replace maintenance supplies

Maintain exterior of home

- Clear debris
- Wash exterior surfaces
- Change storm doors/windows
- Store outdoor furniture and equipment
- Repair outdoor furniture, equipment, and fixtures
- Tend to plants
- Keep walks/drives clear of snow and ice
- Maintain/repair exterior structure
- Other

Clean/repair fabric items (clothes, linens, etc.)

- Store dirty fabrics
- Wash fabrics
- Store clean fabrics
- Repair/mend fabrics
- Store supplies after use

Accident/emergency Procedures

- Identify accident/emergency
- Respond to accident/emergency

Keep living quarters

- Vacuum
- Dust/sweep
- Wash surfaces
- Clean/polish wood surfaces
- Defrost refrigerator/freezer
- Keep rooms tidy
- Tidy storage areas
- Store cleaning supplies
- Removes waste/garbage
- Replace cleaning materials
- Other

Seasonal changes

- Respond to temperature changes
- Respond to chore changes

Budget money

- Obtain income on schedule
- Cash check(s)
- Use bank services
- Allocate funds appropriately

(continued)

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table 3 (continued)

Disruption of routines	Maintain food stock	Home safety
• Identify disruption in routine • Cope with the disruption	• Inventory food stock • Purchase food as required • Store food	• Prevent fires • Prevent accidents • Other
Pay bills	Household supply depletion	Prepare and serve meals
• Pay fixed bills • Purchase necessities • Deposit savings • Pay for recreation • Other	• Identify depleted materials • Respond to such depletions	• Prepare menus • Prepare meals • Serve meals • Preserve and store leftovers • Other
Interaction with community members	Inappropriate conduct of community members	The law
• Perform required interactions • Refrain from interacting inappropriately • Observe demeanor constraints • Observe conversational constraints • Appropriate body language •	• Identify inappropriate conduct of others • Respond to such conduct	• Refrain from illegal acts • Observe the laws • Exercise rights if arrested
Equipment breakdown	Sudden changes in weather	Civic responsibility
• Identify broken equipment • Respond to the breakdown	• Observe sudden weather change • Respond to the required changes in heating • Respond to the required changes in building structure (e.g., screens, storm windows) • Respond to required changes in clothing	• Register for duties • Carry out responsibilities • Volunteer for community services
		Unexpected depletion of funds
		• Identify sudden depletion of funds • Respond to the fund depletion

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table 4

Tasks in the Domain of Leisure (Dever & Knapczyk, 1988)

Find new leisure activities

- Identify potential new activity
- Assess desirability of new activity
- Assess ability to pay required costs

Acquire sills for leisure activities

- Acquire skills through formal lessons
- Acquire skills informally

Perform leisure activity

- Establish activity schedule
- Prepare activity peripherals
- Engage in activities
- Follow usage patterns
- Other

Changes in leisure routine

- Observe schedule changes
- Adapt to schedule changes

Maintain leisure equipment

- Replenish depleted supplies
- Replace worn/defective equipment and materials
- Make repairs and adjustments to equipment
- Store leisure equipment
- Clean clothing /equipment

Leisure safety procedures

- Follow safety rules for leisure activities
- Follow safety rules for leisure facility
- Use equipment appropriately
- Store equipment
- Keep leisure location/equipment free of hazards
- Use safety equipment
- Refrain from entering unsafe areas
- Follow directions of officials
- Observe moderation with alcohol and other substances

Accident/emergency procedures

- Identify accident/emergency
- Report accident/emergency
- Follow accident/emergency action procedures
- Cancel or postpone activity
- Other

Interaction with others

- Perform required actions
- Refrain from inappropriate interactions
- Observe demeanor constraints
- Observe conversational constraints
- Appropriate body language

Inappropriate interaction of others

- Identify inappropriate conduct of others
- Respond to inappropriate conduct of others

(continued)

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table 4 (continued)

Equipment breakdowns and material depletions
• Observe the problem • Adapt to the problem

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table 5

Tasks in the Domain of Travel (Dever & Knapczyk, 1988)

Mental maps of frequented buildings	Follow usage procedures for conveyances	Accident/emergency procedures
• Travel to locations in living quarters (e.g., kitchen, etc.) • Travel to locations within school/work sites • Travel to locations within community sites • Travel to locations at leisure sites	• Ride in/on conveyances (e.g., bike, car) • Operate private conveyances • Ride in public conveyances	• Identify accident/emergency • Report accident/emergency • Follow accident/emergency action procedures • Find alternative travel • Other
Mental maps of the community	Preparation for travel	Interaction with others
• Travel to stores • Travel to medical assistance • Travel to restaurants • Travel to Church • Travel to leisure locations • Travel to bank • Travel to school/work • Travel to center city • Travel between locations	• Select destination • Select transportation • Select departure time • Select appropriate clothing • Obtain funds if necessary • Other	• Perform required interactions • Refrain from inappropriate interactions • Observe demeanor constraints • Observe conversation constraints • Exhibit appropriate body language
	Travel safety procedures	Being lost
	• Follow pedestrian safety rules • Follow bus safety rules • Follow automobile safety rules • Other	• Identify when lost • Respond to being lost

(continued)

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table 5 (continued)

Schedule changes	Inappropriate conduct of others	Equipment breakdowns during travel
• Adjust when conveyance schedule changes • Adjust when personal routine changes • Inform others when unavoidable problems cause schedule changes (e.g., bad weather, equipment breakdown, etc.) • Other	• Identify inappropriate conduct of others • Respond to inappropriate conduct of others	• Identify equipment breakdown • Respond to equipment breakdown

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table 6

Skill Categories and Items in the Assessment of Motor and Process Skills (Fisher 2006a, 2006b)

Motor skills		Process skills	
Body position • Stabilizes • Aligns • Positions	Moving self and objects • Moves • Lifts • Walks • Transports • Calibrates • Flows	Sustaining performance • Paces • Attends • Heeds	Organizing space and objects • Searches/Locates • Gathers • Organizes • Restores • Navigates
Obtaining & holding objects • Reaches • Bends • Grips • Manipulates • Coordinates	Sustaining performance • Endures • Paces	Applying knowledge • Chooses • Uses • Handles • Inquires	Adapting performance • Notices/Responds • Adjusts • Accommodates • Benefits
		Temporal organization • Initiates • Continues • Sequences • Terminates	

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table 7

The Functional Independence Measure (FIM) Items and Levels of Scoring (USD_{MR}, 1997)

FIM™ motor items	FIM™ cognitive items	Levels of scoring
<u>Self-care</u>	<u>Communication</u>	7 – complete independence
• Eating	• Comprehension	6 – Modified independence (use of an assistive device)
• Grooming	• Expression	
• Bathing		
• Dressing upper body	<u>Social cognition</u>	5 – Supervision
• Dressing lower body	• Social interest	4 – Minimal assistance (client performs $\geq$ 75% of the task)
	• Problem solving	
• Toileting	• Memory	
<u>Sphincter control</u>		3 – Moderate assistance (client performs 50-74%)
• Bladder management		2 – Maximal assistance (client performs 25-49%)
• Bowel management		1 – Total assistance (client performs 0-25%)
<u>Transfer</u>		
• Bed, chair, wheelchair		
• Toilet		
• Tub, shower		
<u>Locomotion</u>		
• Walk/wheelchair		
• Stairs		

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table 8

Items within the Functional Autonomy Measurement System (SMAF)

Activities of daily living	Communication	Social functioning*
• Eating • Bathing • Dressing • Grooming • Urinary continence • Fecal continence • Using the bathroom	• Seeing • Hearing • Talking	• Occupies his/her free time • Maintains/creates significant relationships • Uses resources of his/her environment • Acts appropriately in relationships with others • Accomplish significant roles appropriate to own situation • Expresses desires, ideas, opinions and limitations
Mobility	**IADL**	
• Transfers • Walking inside • Walking outside • Putting on prosthesis and/or orthosis • Moving in a wheelchair • Using the stairs	• Cleaning the house • Preparing meals • Shopping • Doing the laundry • Using the phone • Using public transportation • Taking medication • Managing the budget	
	Mental functions	
	• Memory • Orientation • Understanding • Judgment	

* Items within social functioning subscale are still being developed

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table 9

Summary Scores for the Katz Index of ADL

Summary grade	Definition of summary grade
A	Independent in feeding, continence, transferring, going to the toilet, & bathing
B	Independent in all but one of these functions
C	Independent in all but bathing and one additional function
D	Independent in all but bathing dressing, and one additional function
E	Independent in all but bathing, dressing, going to the toilet, and one additional function
F	Independent in all but bathing, dressing, going to the toilet, transferring, and one additional function
G	Dependent in all six functions
Other	Dependent in at least two functions, but not classifiable as C, D, E, or F

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table 10

Sample Item from the Klein-Bell ADL Scale

C. Socks
8. Grasp sock
9. Reach sock to R foot
10. Reach sock to L foot
11. Pull sock over L toes
12. Pull sock over R toes
13. Pull sock over R foot with heel to heel
14. Pull sock over L foot with heel to heel
15. Pull sock up to full extension of R leg
16. Pull sock up to full extension of L leg

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table 11

Global Comparison of ADL Assessment Properties

Assessment	PADL items	IADL items	Type of data	Ceiling/floor effects
AMPS	Yes	Yes	Interval	No
FIM	Yes	No	Ordinal	Yes
SMAF	Yes	Yes	Ordinal	Unknown
KATZ	Yes	No	Ordinal	Yes
Klein-Bell	Yes	No	Weighted	Yes

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table 12

Tasks within the Community Living Skills Assessment Inventory (Switzky, Rotatori, & Cohen, 1978)

<u>Dressing and undressing</u>

- Select clothing
- Put clothing on
- Remove clothing
- Put on outerwear
- Take of outerwear

<u>Personal hygiene & grooming</u>

- Bathe/shower
- Brush teeth
- Care for hair
- Wash up
- Use the toilet
- Care for nails
- Shave

<u>Eating</u>

- Use spoon, fork, napkin
- Drink from glass
- Pass food
- Pour liquid

<u>Housekeeping</u>

- Make bed
- Vacuum
- Dust
- Clean tub/shower, toilet sink, etc.
- Mop, sweep, wax floor

<u>Caring for clothing</u>

- Store dirty laundry in appropriate place
- Sort clothing
- Load washing machine
- Operate washing machine
- Fold clothes
- Iron clothes
- Hang clothes
- Load and operate dryer
- Repair torn clothing
- Care for shoes
- Sew buttons on

<u>Functional/adaptive equipment</u>

- Manage leg brace (e.g., take on/off, secure the brace)
- Manage eye glasses
- Manage hearing aid
- Manage wrist watch
- Wheelchair (manage brakes, footrests, propel self)
- Transfer into/out of wheelchair

<u>Functional/adaptive equipment (continued)</u>

- Wheelchair mobility (e.g., up/down ramps, use elevator, operate light switches, up/down stairs)
- Utilize crutches
- Use adaptive eating utensils

<u>Preparing food</u>

- Prepare cold cut sandwich
- Cook food (boil, fry, bake, grill, broil)
- Load dishwasher
- Run garbage disposal
- Operate toaster
- Set table
- Clear table
- Hand wash dishes
- Hand dry dishes

<u>Self-medication</u>

- Know when to take meds
- Know what pills to take
- Know number of pills to take
- Open pill box
- Put pill(s) in hand
- Put pills in mouth
- Swallow pills
- Get drink of water

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table 13

Scoring Criteria for the Community Living Skills Assessment Inventory (Switzky, Rotatori, & Cohen, 1978).

Score	Criteria
6	Independent, perfect performance: Individual accomplishes the particular item correctly and completely by himself. If the rater must demonstrate the task, the person's skill cannot be scored as independent.
5	Independent, imperfect performance: Individual accomplishes the particular item completely by himself, but makes some error(s) (e.g., puts shoes on wrong feet, shirt on inside out, etc.)
4	Supervised: Individual can accomplish the particular items my himself, but must be supervised (e.g., given directions, given items), and/or a demonstration of the task must be provided.
3	Assisted, partially: Individual accomplishes more than half of the task by himself and is assisted the remainder of the time.
2	Assisted, primarily: Individual accomplishes less than half of the task by himself and is assisted the remainder of the time.
1	Cooperative: Individual merely implements some of the task by positioning himself in a manner appropriate to the item.
0	No participation: Individual remains completely passive or may actually offer resistance to the task performed by the rater.
NA	Not applicable: Due to the individual's biological or physical development the item is not applicable (e.g., wheelchair items for one that is ambulatory).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table 14

Items on the St. Louis Inventory of Community Living Skills (Evenson & Boyd, 1993)

1. Personal hygiene – bathing, showering, dental care, and general cleanliness throughout the day.

2. Grooming – hair and hand care, shaving/makeup, and general appearance.

3. Dress skills – dressing, selecting appropriate garments, keeping clothes neat throughout the day.

4. Self-Care – bed making, cleaning, organizing possessions, etc.

5. Communications – ability to ask for help when needed, make wishes known, express feelings in appropriate ways, engage in meaningful conversations, etc.

6. Safety – proper use of cigarettes, matches, tools, writing materials, glasses, etc.

7. Handling time – ability to keep appointments, return when due, participate in activities when scheduled, arrive to work on time, etc.

8. Handling money – responsible spending habits, budgeting ability, and keeping money safe.

9. Leisure activities – some use of variety of activities such as newspapers, other reading, TV, structured activities, hobbies, sewing, card playing, etc.

10. Clothing maintenance – doing laundry, ironing sewing buttons, keeping clothes intact, using hangers, etc.

11. Meal preparation – cooking, shopping, serving, using a microwave, etc.

12. Sexuality – safe sex

13. Use of resources – use of telephone, transportation, cafeteria/restaurant, shopping, and other community resources.

14. Problem solving – ability to seek help when needed, basic skills in decision making, ability to ask questions, etc.

15. Health practices – proper nutrition, weight control, exercise, handling medications, etc.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table 15

Number of Participants by Diagnostic Group and Global Functional Level

	Global functional level			
Diagnostic category	Independent	Minimal assistance	Moderate assistance	Total
Well	1226	0	0	1226
At risk	424	21	0	445
Frail older adult	0	61	6	67
Mild learning disability	42	52	32	126
Neurologic developmental	57	315	636	1008
Mental retardation	44	621	1015	1680
Other neurological	1055	3258	4454	8767
Hemispheric stroke	777	3145	4879	8801
Musculoskeletal	2598	3003	1780	7381
Medical	705	1009	851	2565
Psychiatric	1580	3998	2978	8556
Dementia	99	681	1708	2488
Other memory	27	73	90	190
Other	1824	7442	11900	21166
Total	10458	23679	30329	64466

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table 16

Mean ADL Motor and ADL Process Ability by Global Functional Level

	Global functional level		
	Independent (n = 10,458)	Minimal assistance (n = 23,679)	Maximal assistance (n = 30,329)
Age (years)			
M	55.1	57.6	61.7
SD	17.9	20.2	20.6
ADL motor ability (logits)			
M	1.83	1.30	0.64
SD	0.69	0.76	0.98
ADL process ability (logits)			
M	1.45	0.99	0.35
SD	0.51	0.48	0.96

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Table 17

ROC Curve Analysis: Using ADL Motor or ADL Process Ability as an Indicator of Community Independence

	Independent vs. assistance cutoff measures (logits)		Independent vs. assistance cutoff measures (logits)		Other vs. maximum assistance cutoff measures (logits)	
	Motor	Process	Motor	Process	Motor	Process
	2.0	1.0	1.5	1.1	1.0	0.7
Area under the curve	0.78	0.84	0.78	0.84	0.74	0.82
Sensitivity*	0.40	0.81	0.67	0.78	0.70	0.77
Specificity**	0.87	0.70	0.72	0.73	0.66	0.69

* Estimated true positive rate
** Estimated true negative rate

Table 18

*Area Under the Receiver Operating Characteristic (ROC) Curve by Decision Condition and Diagnostic Category**

		Area under the ROC curve			
Diagnostic category	*n*	Independent vs. in need of assistance		Other vs. maximum assistance	
		Motor	Process	Motor	Process
Neurologic developmental	1,008	0.73	0.85	0.72	0.76
Mental retardation	1,680	0.75	0.83	0.67	0.75
Other neurological	8,767	0.79	0.83	0.76	0.81
Hemispheric stroke	8,801	0.82	0.82	0.77	0.80
Musculoskeletal	7,381	0.82	0.77	0.85	0.79
Medical	2,565	0.85	0.81	0.83	0.81
Psychiatric	8,556	0.68	0.77	0.66	0.76
Dementia	2,488	0.78	0.92	0.70	0.84
Other memory	190	0.85	0.91	0.74	0.83
Other	21,166	0.79	0.83	0.73	0.81
Entire sample	64,466	0.78	0.84	0.74	0.82

* Well, mild learning disability, and frail diagnostic categories were not analyzed due to small sample sizes and invariant functional level ratings

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

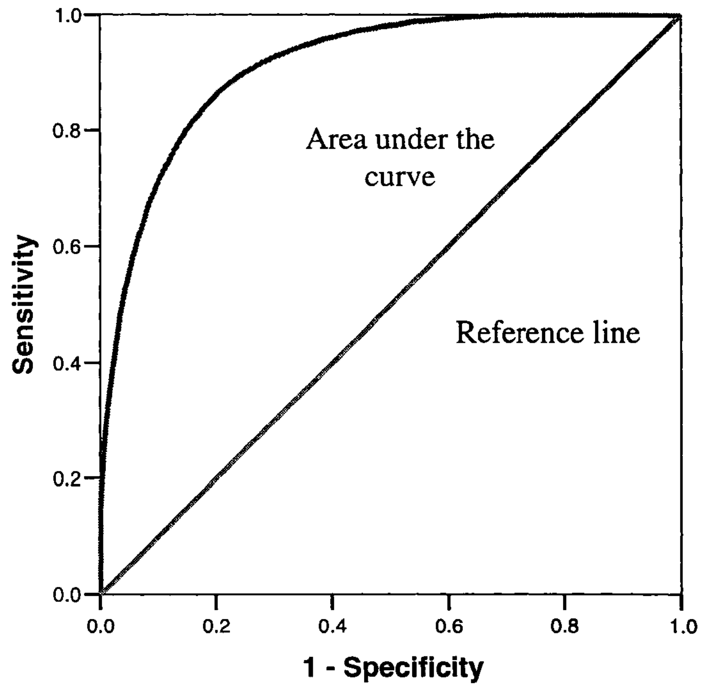


Figure 1. Example of a receiver operating characteristic (ROC) curve

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

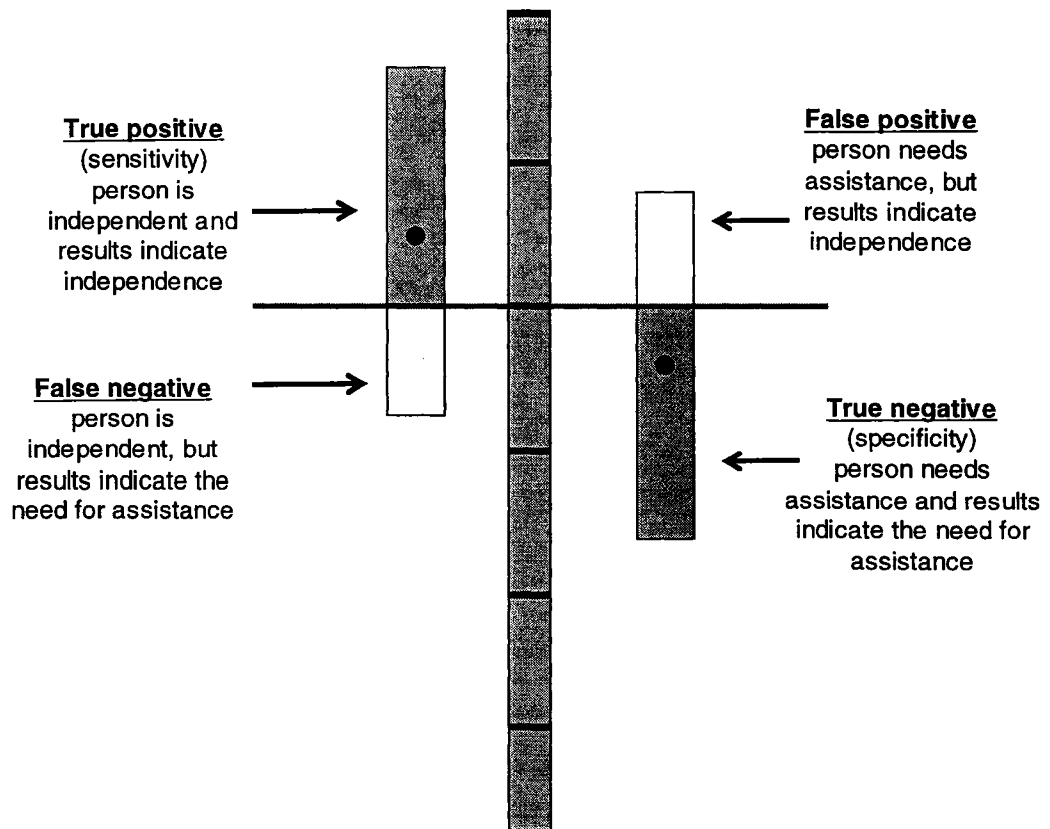


Figure 2. Identification and definition of the four possible decisions within the current project

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

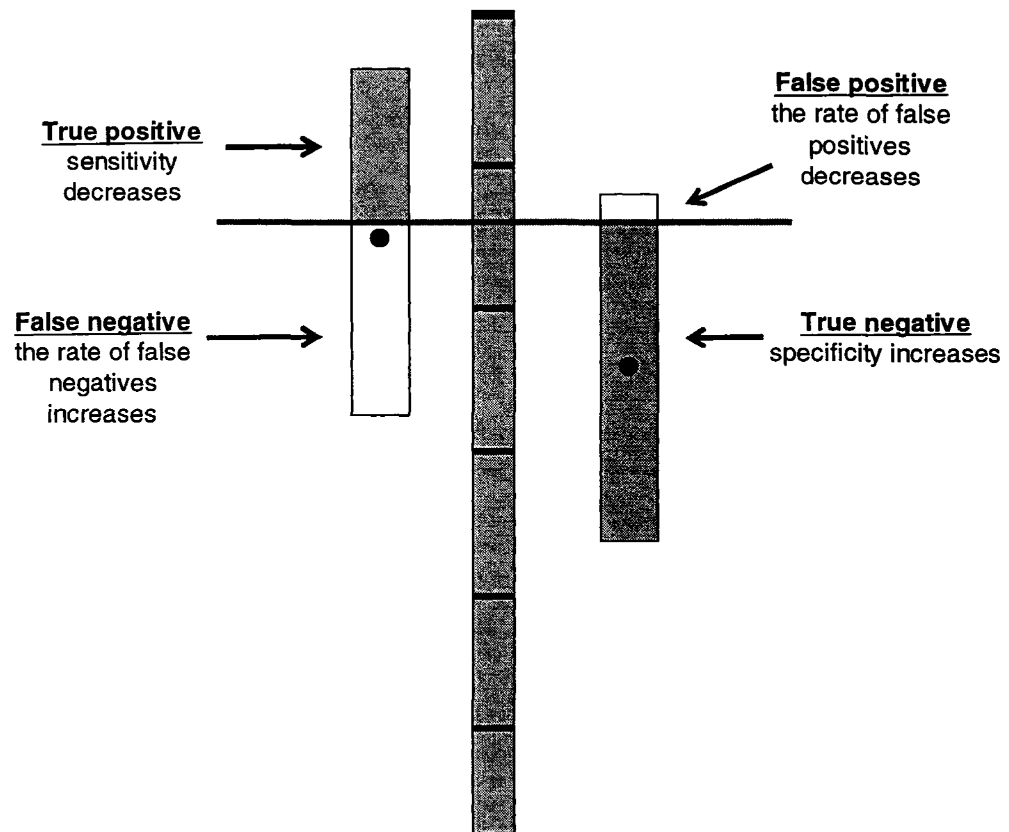


Figure 3. Impact of raising the cutoff measure or decision threshold (compare to Figure 4)

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

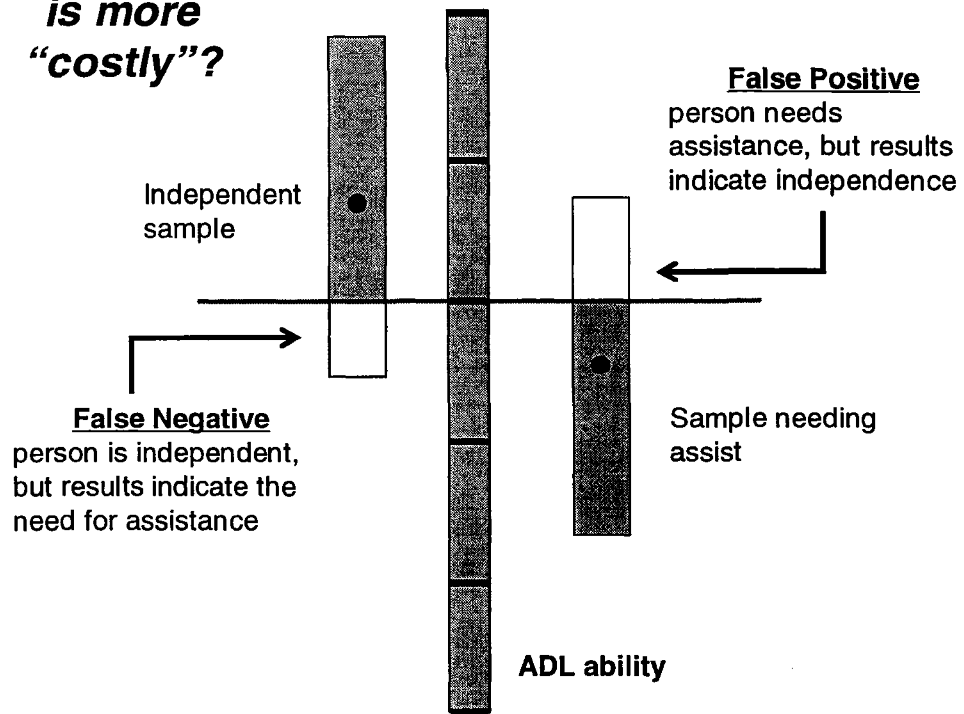


Figure 4. Consideration of the relative costs of possible errors

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

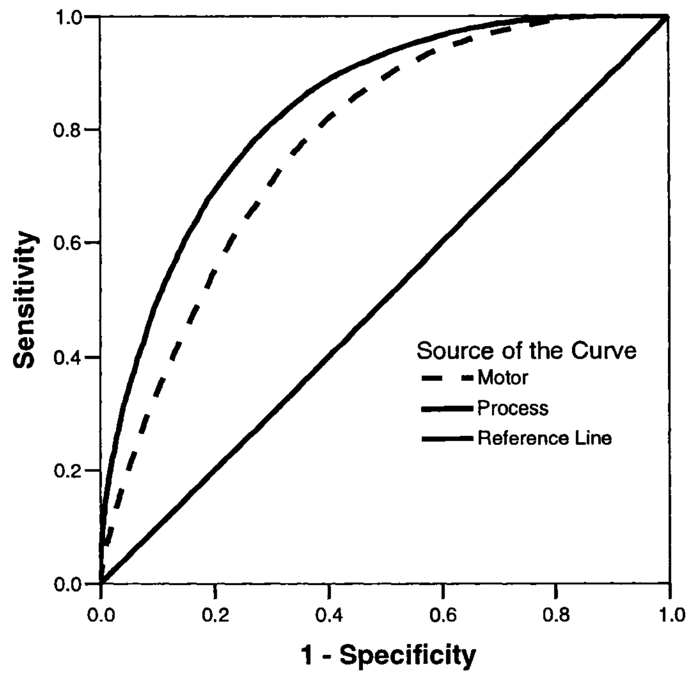


Figure 5. ROC curve using ADL motor or ADL process ability to categorize individuals as either independent in the community or in need of assistance to live in the community

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

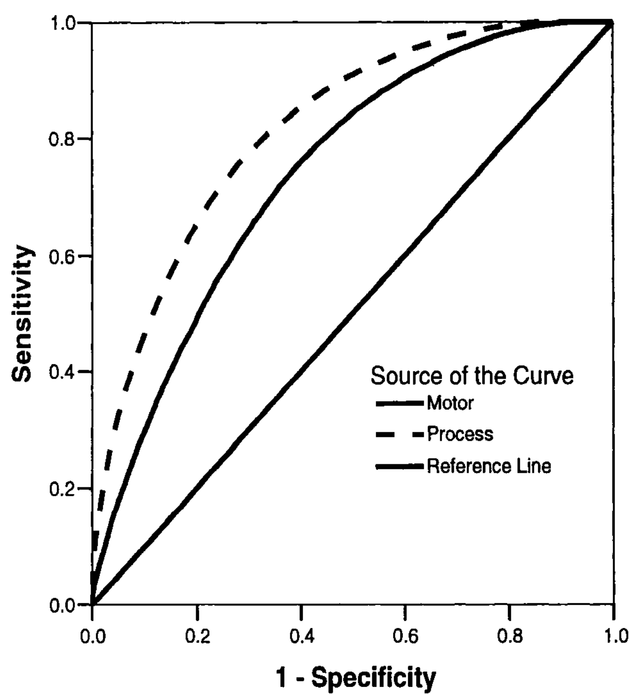


Figure 6. ROC curve using ADL motor or ADL process ability to categorize individuals either as "other" or in need of moderate to maximal assistance to live in the community

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

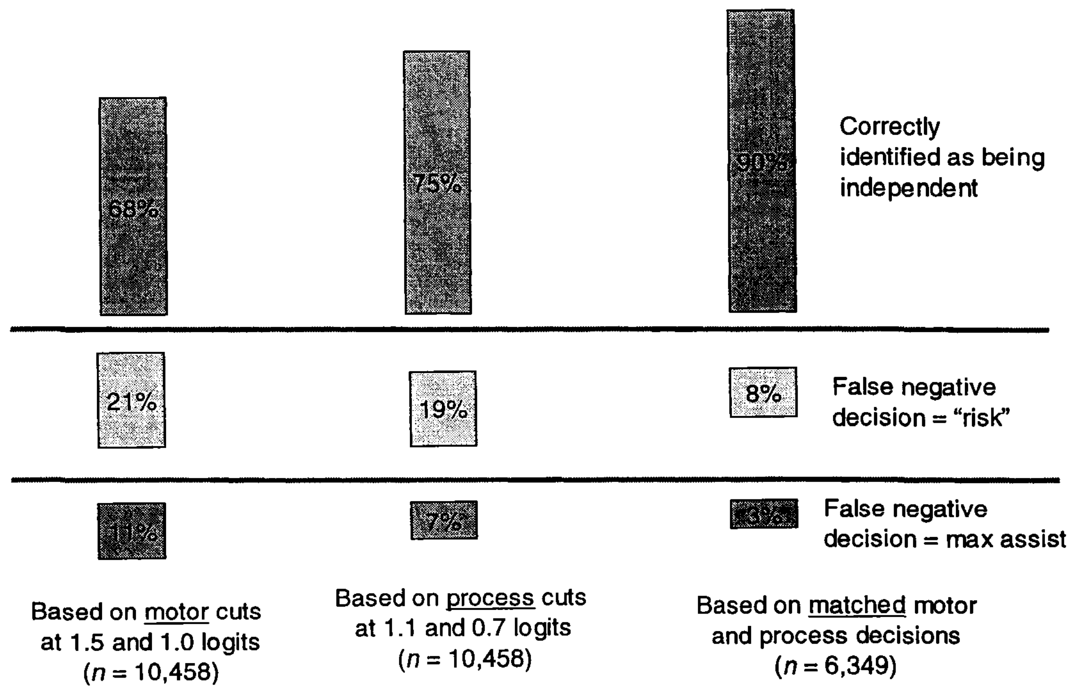


Figure 7. Accuracy of correctly categorizing individuals who are independent in the community

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

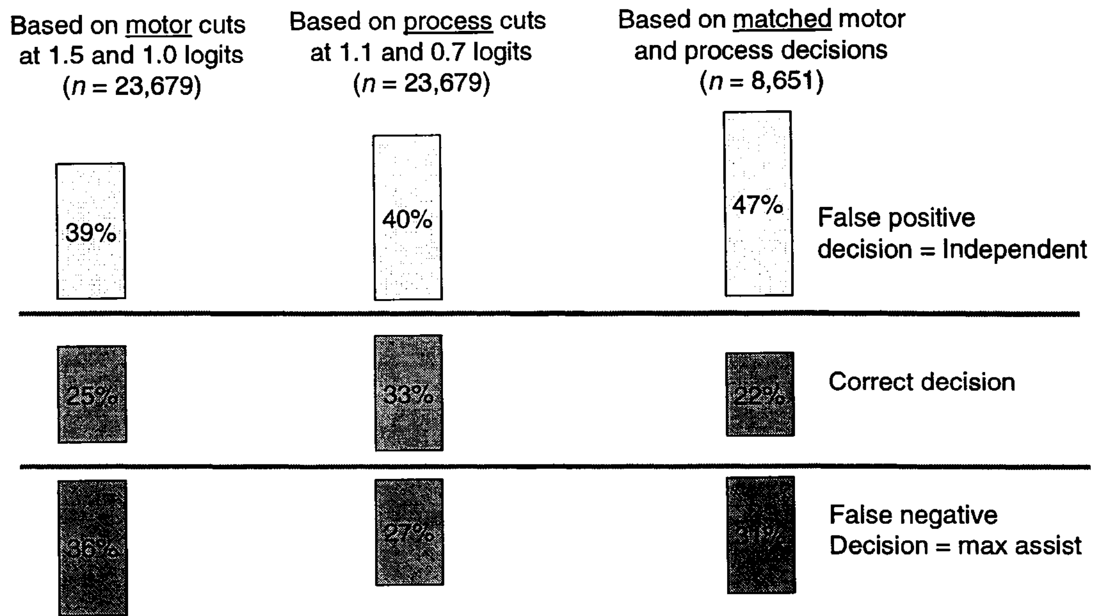


Figure 8. Accuracy of correctly categorizing individuals who are in need of minimal assistance to live in the community

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

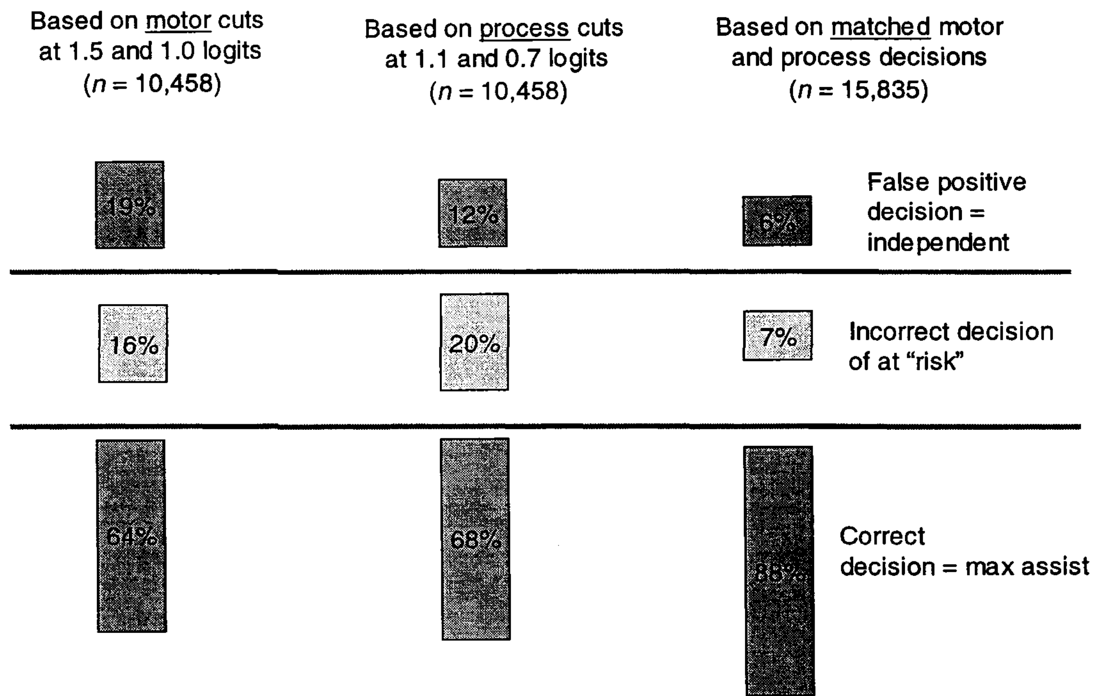


Figure 9. Accuracy of correctly categorizing individuals who are in need of moderate to maximal assistance to live in the community

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Appendix A

Definition of Occupational Therapy Terms

Activity: What one observers (action or doing) as one performs a task (Fisher, 2006c).

Activities of daily living (ADL): In the context of this paper, globally refers to both personal and instrumental activities of daily living.

Basic or personal activities of daily living (PADL): An area of occupation, oriented toward caring for one's own body (e.g., bathing, eating, dressing, sleep, sexual activity, toilet hygiene) (AOTA, 2002).

Independent living skills: The ability to perform all daily life tasks as needed for independent living in the community, including the ability to direct or procure someone else to perform the selected tasks, as needed or desired by the individual.

Instrumental activities of daily living (IADL): An area of occupation, oriented toward interacting with the environment (e.g., caring for others, caring for pets, home maintenance, cooking, safety procedures, shopping) (AOTA, 2002).

Community living skills: See independent living skills.

Level of independence (or assistance): The level of support or assistance needed for living in the community. Includes any assistance needed, whether or not it is independently procured by the individual.

Occupation: A series of task-related and goal-directed actions in which one is engaged (Fisher, 2006c).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Occupational performance: Carrying out daily life tasks in the assorted areas of occupation (i.e., PADL, IADL, education, work, play, leisure, and social participation). Also refers to the accomplishment of the selected task, resulting from the dynamic interaction of the person, context, and activity (AOTA, 2002).

Task: What the person will do or has done once the performance is completed (Fisher, 2006c).

Task performance: The performance of a daily life task, with or without engagement (Fisher, 2006c).

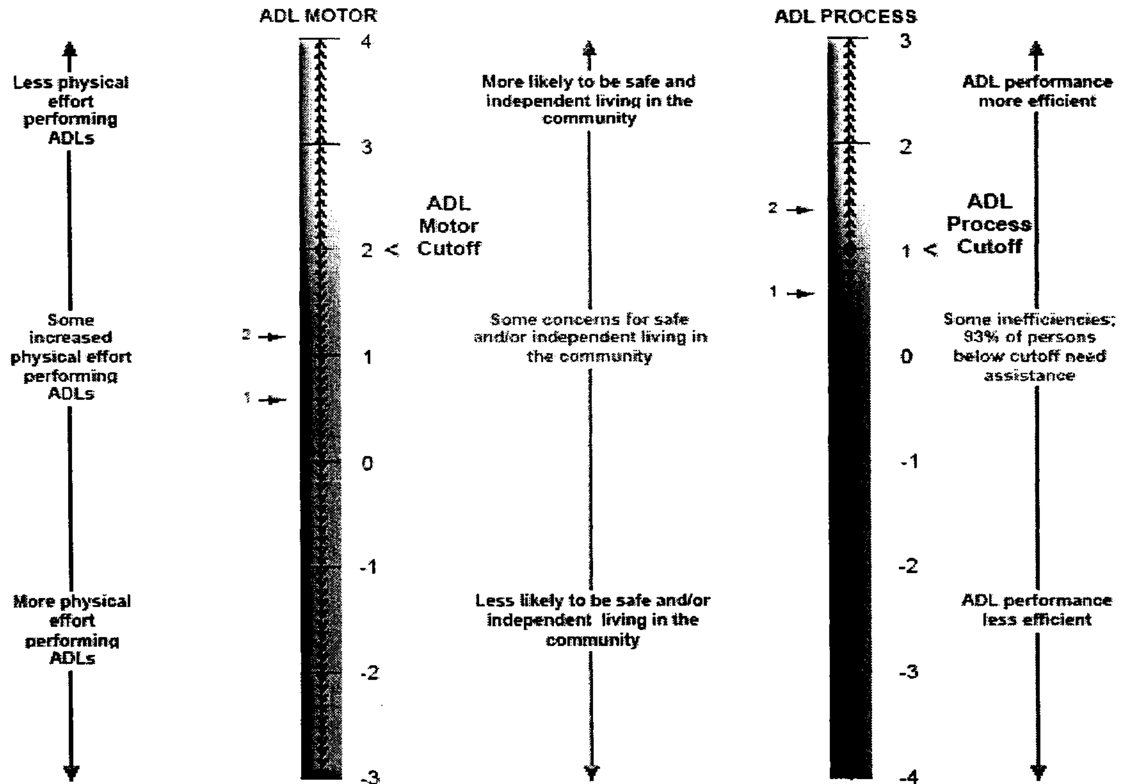
Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Appendix B

Sample AMPS Graphic Report

ASSESSMENT OF MOTOR AND PROCESS SKILLS (AMPS) GRAPHIC REPORT

Client:	Sample R		DATE	MOTOR	PROCESS
Occupational therapist:	Brenda K. Merritt	Evaluation 1	04/28/2006	0.66	0.60
	Brenda K. Merritt	Evaluation 2	05/30/2006	1.21	1.46



The numbers on the ADL motor and ADL process scales are units of ADL ability (logits). The results are reported as ADL motor and ADL process measures plotted in relation to the AMPS scale cutoffs. Measures below the cutoffs indicate that there was diminished quality or effectiveness of performance of instrumental and/or personal activities of daily living (ADL). See the AMPS Narrative Report for further information regarding the interpretation of a single AMPS evaluation.

Graphic Report | Printed: 07/31/2006 | Page 1

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Appendix C

Definition of Terms Related to Receiver Operating Characteristic (ROC) Curves

Decision threshold: The cutoff measure; a measure above the cutoff indicates a positive result and a measure below the cutoff value indicates a negative result.

False negative: The test incorrectly identifies an individual as not having the condition when the condition is present. In the current project, a false negative error occurs when an individual is identified as needing assistance to live within the community, when in fact he/she does not need assistance.

False positive: The test incorrectly identifies an individual as having the condition when the condition is not present. In the current project, a false positive error occurs when an individual is identified as being independent within the community, when in fact he/she needs assistance.

Sensitivity: The true positive rate; how the test performs among those with the condition. The ability of the test to detect the condition when the condition is present. Within the current project, the true positive rate is the rate of those who are correctly identified as being independent within the community.

Specificity: The true negative rate; how the test performs among those without the condition. The ability of the test to exclude the condition in those without the condition. In the current project, the true negative rate is the rate of those who are correctly identified as needing assistance to live in the community.

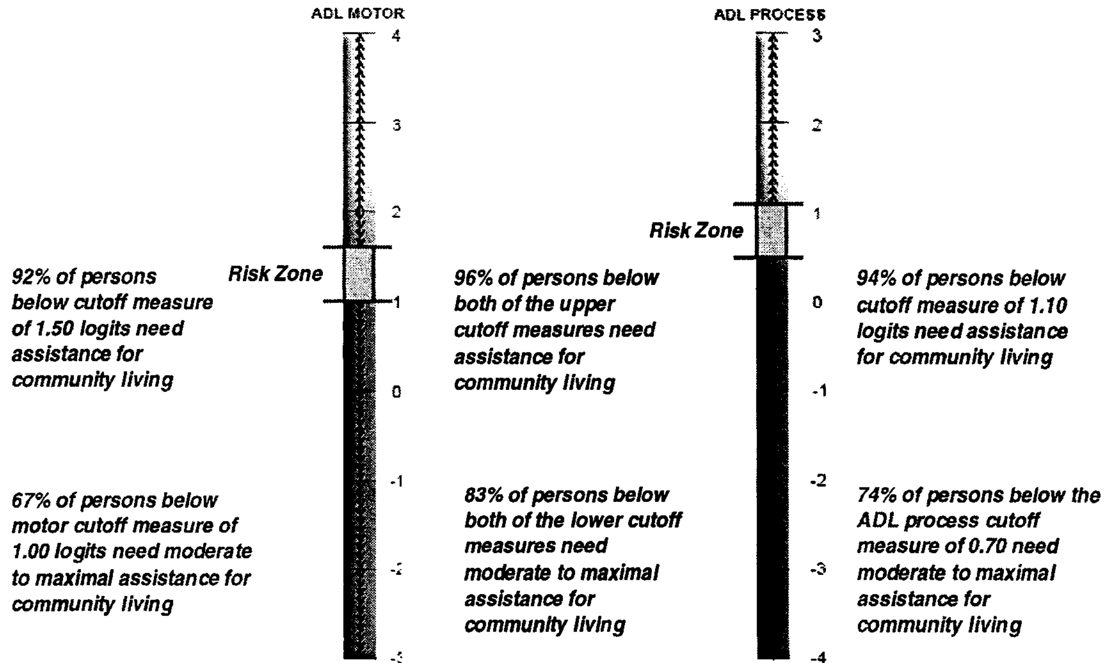
True positive rate: See sensitivity

True negative rate: See specificity

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Appendix D

Proposed Changes to the AMPS Graphic Report*



* For graphical simplicity, other aspects of the Graphic Report were eliminated in this figure, but would be included in the final version

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.