TASK PERFORMANCE OF OLDER WOMEN WITH MAJOR DEPRESSION

by

Razan Hamed

BS, University of Jordan, 2003

MS, University of Pittsburgh, 2005

Submitted to the Graduate Faculty of

the School of Rehabilitation Sciences in partial fulfillment

of the requirements for the degree

of Doctor of Philosophy in Rehabilitation Sciences

University of Pittsburgh

2008

UNIVERSITY OF PITTSBURGH

SCHOOL OF HEALTH AND REHABILITATION SCIENCES

This dissertation was presented

by

Razan Hamed

It was defended on

April 25, 2008

and approved by

Richard Morycz, PhD, Department of Psychiatry

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

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

Denise Chisholm, PhD, OTR/L, Department of Occupational Therapy

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

TASK PERFORMANCE OF OLDER WOMEN WITH MAJOR DEPRESSION

Razan Hamed

University of Pittsburgh, 2008

Depression is a serious mental health illness and one of the most burdensome diseases to society. In fact, it is the most costly disease among mental illnesses in the United States (Murray & Lopez, 1997). The assessment of task performance is an important contribution provided by occupational therapists in the care of patients with depression. Such assessment can detect various difficulties experienced by patients. In older patients these difficulties are more complicated due to aging. Examples of difficulties may be failing to dress oneself, inability to take medication on time, and needing assistance in preparing a meal. Factors that underlie these difficulties might be age-related factors (e.g., cognitive and physical impairments), environment-related factors (e.g., novelty of the context in which the task was performed), or depression-related factors (e.g., negative symptoms associated with depression [e.g., loss of motivation]). The purpose of this dissertation is to determine difficulties in task performance experienced by older women with late life depression.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

My doctoral dissertation would have never been completed without the help of many people.

First, my deepest appreciation goes to Margo B. Holm, PhD, OTR/L, my advisor, not only for her academic guidance, support, and patience, but also for the wonderful friend she has been to me during my studies. The skills I learned from Dr. Holm are invaluable. Also, I would like to thank the members of my committee, Joan C. Rogers, PhD, OTR/L, for sharing her abundant knowledge and expertise and taking the time to patiently nurture my professional development; Dr. Richard Morycz, PhD, for his expertise in psychiatry and for guiding me to see the "big picture" of my research; Dr. Elizabeth Skidmore, PhD, OTR/L, for sharing her professional perspectives and expertise and for the verbal and nonverbal support she has given me; and Denise Chisholm, PhD, OTR/L for her significant academic and professional contributions throughout my studies.

Also, I would like to thank my colleagues Min-Mei Shih, PhD, OTR, and Ketki Raina, PhD, OTR/L, and all my friends for their immeasurable assistance and collaboration throughout my doctoral studies.

Finally, I would like to thank the dearest to my heart, my mother, Fawzieh, and father Taiseer, and all my siblings who continued to send their love and support across the oceans to help me accomplish my goals.

Thank you all for helping me be where I am today.

1.0 INTRODUCTION

Major depression is one of the most prevalent psychiatric disorders in the United States older adult population (Kessler et al., 2005; National Institute of Mental Health, 2007). Up to 15% to 20% of older adults have significant depressive symptoms (Gallo & Lebowitz, 1999; National Institute of Health, 1992), and 1% to 5% meet the DSM-IV criteria for clinical major depression (Gurland, Cross, & Katz, 1996); National Institute of Mental Health, 2007). Depression affects both men and women of any age, but it is twice as common in middle-aged and older women as in men (Desai & Jann, 2001). Persistent sadness, feelings of hopelessness and guilt, loss of motivation and pleasure, fatigue, and psychomotor retardation are examples of the DSM-IV-TR criteria for major depression (American Psychiatric Association, 2000). Not only are these symptoms of major depression associated with increased medical costs and healthcare utilization (Druss, Rohrbaugh, & Rosenheck, 1999; Unützer et al., 1997) but they are also associated with poor quality of life and significant functional disability in daily activities (Alexopoulos et al., 1996; Berkman et al., 1986; Bruce, 1999; Callahan et al., 1998; Geerlings, Beekman, Deeg, Twisk, & Van Tulburg, 2001; Rogers & Holm, 2000; Steffens, Hays, Krishnan, 1999; Stuck et al., 1999). In 1990, major depression was the leading cause of functional disability, and it is expected to be the second by 2020 (Lopez & Murray, 1998). Furthermore, research consistently shows that women exhibit more functional disability as they are often responsible for home management (Fried & Guralnik, 1997). Additionally, women are known to live longer than men

(Fried & Guralnik, 1997), and therefore, they often live with disability in daily activities for a longer time than men.

Functional disability can be observed in functional mobility (FM) such as indoor and outdoor walking, stair use, and functional transfers, in activities of daily living (ADL) such as dressing, bathing, and grooming, and in instrumental activities of daily living (IADL) such as meal preparation, shopping, and money management (Nichols, 1974; Spector, 1996). Functional disability, as a consequence of major depression, in these domains can be significantly debilitating to people regardless of their age (National Institute of Mental Health, 2007). However, for older adults with major depression, this disability is confounded by increased age. That is, older adults often exhibit age-related cognitive (e.g., memory loss and slow processing speed) and physical impairments (loss of muscle strength and endurance) that interfere with independent performance of daily activities (Alexopoulos et al., 1996; Steffens, Hays, Krishnan, 1999). Moreover, the depressive symptoms experienced by older adults such as depressed mood (Alexopoulos et al., 1996; Forsell, Jorn, Winblad, 1994), apathy, psychomotor retardation, loss of motivation (Alexopoulos et al., 1996), lack of energy, and difficulties in thinking and concentration (Rogers & Holm, 2000) are significantly associated with limitations in ADL and IADL. Indeed, research shows that older adults with major depression exhibit limitations in ADL and IADL significantly greater than those exhibited by older adults with physical illnesses such as diabetes and back problems (Wells et al., 1989). These limitations were also shown to persist over time (Callahan et al., 1994). In addition, research further shows that the severity of depression is associated with the extent of limitations in daily activities (Ormel et al., 1998). The burden caused by depression may explain the fact that older adults tend to worry more about the risk of losing independence in ADL and IADL than about the depression itself (Lachs et al.,

1990). Unfortunately, the fact that older adults with major depression often exhibit both age-related and affective impairments, may contribute to the misconception that major depression is a natural consequence of aging (Schulberg et al., 1995). As a consequence, major depression has become widely underrecognized and undertreated in older adults (Desai & Jann, 2001; Kessler et al., 2005; Schulberg et al., 1995; Sheehan, 2004; Wells et al., 1989).

Hence, given the significance of ADL and IADL for older adults' quality of life (Bruce, Seeman, Merril, & Blazer, 1994), and the significance of independence in these activities as one measure of successful functioning (Rogers et al., 2001) understanding the disability in ADL and IADL for older adults with major depression is critical. Moreover, as women are more vulnerable to major depression and its consequential disability in ADL and IADL, it is important to understand disability in ADL and IADL in older women to help minimize dependence and maximize independence in daily activities. This, in turn, may improve the quality of life for this vulnerable population.

Despite the rising number of people diagnosed with depression (Bartels & Smyers, 2002) and the evident disability associated with major depression in older adults, especially women, there have been no studies that specifically investigated functional disability in FM, ADL, and IADL for older women with major depression. Therefore, in this dissertation we investigate difficulties in task performance for older women with major depression. The specific aims for the study are to:

(1) Locate and evaluate evidence of the influence of the environment and time on task performance for older women with major depression

(2) Compare the influence of two different environments on task performance for older women with major depression.

(3) Examine the effect of time on task performance for older women with major depression.

Chapter 2, 3, and 4 of this dissertation present our studies for aims 1, 2, and 3 respectively.

2.0 THE INFLUENCE OF THE ENVIRONMENT AND TIME ON TASK PERFORMANCE IN OLDER ADULTS WITH MAJOR DEPRESSION: TWO SYSTEMATIC REVIEWS

2.1 BACKGROUND AND SIGNIFICANCE OF THE PROBLEM

The older adult population is rapidly growing in the United States (US) with over 12.4% of the population being older than 65 years (Cristian, 2006). By 2030, 20% of the population will be over the age of 65: almost a twofold increase since 2005 (US Department of Health and Human Services, 2006). This segment of the population exhibits a wide variety of limitations in basic activities of daily living (ADL [also known as personal care activities of daily living {PC}]), and instrumental activities of daily living (IADL) (Dunlop, Hughes, & Manheim, 1997; Fried & Guralink, 1997). Over 27% of older adults report limitations in one or more ADL, and 13.7% reported limitations in IADL (US Department of Health and Human Services, 2006). Remediating these limitations uses over 35% of the US healthcare expenditures annually (Fried, Bradley, Williams, & Tinetti, 2001). Although these limitations are report by both men and women, research consistently shows that women live longer than men, and therefore, they live with these limitations for a longer time (Fried & Guralnik, 1997). Furthermore, as women are usually more responsible for home management tasks than men as part of their role-related responsibilities (Brines, 1994; Marini & Shelton, 1993; Shelton & John, 1996), it is expected that

women will experience more limitations than men in ADL and IADL. Hence, at all times, a larger number of older women will report limitations in ADL and IADL compared to older men.

Several risk factors may lead to the limitations in ADL and IADL in old age including major depression; age-related cognitive; physical, and visual impairments; disease burden; and low frequency of social contacts (Stuck et al., 1999). Among these factors, major depression was considered the most debilitating factor in old age (Blazer, 2003; Stuck et al., 1999). Every year, over 4% of the older adult population experiences the negative symptoms of major depression, with twice the prevalence in women than men (National Institute of Mental Health, 2007). Depressed mood, psychomotor retardation or agitation, feelings of guilt, loss of motivation, apathy, and fatigue, are examples of these negative symptoms (American Psychiatric Association, 2000). These symptoms can be significantly debilitating to the independent performance of ADL and IADL, in that, older adults with major depression may need assistance in initiating, continuing, or completing essential tasks due to these disabling symptoms (Rogers & Holm, 2000). Normal age-related cognitive (e.g., longer reaction time to environmental stimuli) and physical impairments (e.g., loss of strength and endurance) may further add to the debilitating effect of major depression in older adults (Callahan, 2005). Despite the fact that research has consistently shown a significant association between major depression and limitations in ADL and IADL in several older adult samples (Beekman, Deeg, Braam, &Van Tilburg, 1997; Bruce, 1999; Prince, Harwood, Blizard, Thomas, & Mann, 1997; Stuck et al., 1999; Wells et al., 1989) and that older adults with major depression show greater difficulties in daily activities than older adults with physical conditions (Wells et al., 1989), major depression is often underrecognized and is considered a consequence of comorbid diseases or a natural result of becoming older rather than a major risk factor for functional disability (limitations in

ADL and IADL) (Schulberg et al., 1995). As a result, major depression has become a widely underrecognized and undertreated illness in older adults (Desai & Jann, 2001; Kessler et al., 2005; Schulberg et al., 1995; Sheehan, 2004; Wells et al., 1989).

Another risk factor that may lead to limitations in task performance in older adults that has been less investigated in past research as much as other risk factor is the environment (Stuck et al., 1999). The environment where people perform their daily activities can be either natural and familiar, such as the home, or standardized and unfamiliar, such as a clinical settings or facilities (e.g., nursing facilities and/or psychiatric, medical, and rehabilitation hospitals). For simplification, we will refer to "clinical setting" as "clinic" throughout this dissertation. That is, people usually perform daily tasks in their homes (Yerxa & Baum, 1987). However, sometimes healthcare practitioners ask patients to perform certain tasks in the clinic setting for functional assessment purposes (Park, Fisher &Velozo, 1994). Some people find that performing tasks at home is easier than in the clinic, and some find it easier in the clinic than at home. This indicates that both the home and the clinic can be either facilitating or demanding for certain people, or for certain tasks. The facilitative or demanding nature of the environment may affect the final outcome of task performance at home during routine activities or in the clinic during a functional assessment (Christiansen & Baum, 1991; Davidson, 1991; Nichols, 1976). Subsequently, this may affect decisions about the patient's overall functional ability (overall level), the patient's ability to perform tasks within a domain (e.g., IADL), or the patient's ability to perform a specific task within a domain (e.g., meal preparation).

Lawton (1982; 1983; 1986) was the first to discuss the association between the *competence* of the individual (i.e., the ability to successfully perform daily tasks), and *environmental demands* (i.e., challenges imposed by the performance environment). He

explained that high competence is associated with an adaptive behavior to the environmental demands and positive affect, whereas low competence is associated with a maladaptive behavior to the environmental demands, and negative affect. In other words, when the individual's competence is high in comparison to or matches the environmental demands, the individual will exhibit an adaptive behavior (i.e., successfully overcomes the difficulties imposed by the environment) and positive affect (e.g., experiences high self-esteem) (Czaja, Weber, & Nair, 1993). In contrast, when the individual's competence is too low in comparison to, or does not match the environmental demands, the individual will exhibit a maladaptive behavior (i.e., does not perform the task or performs the task only with assistance) and negative affect (e.g., experiences low self-esteem) (Czaja, Weber, & Nair, 1993). For example, a frail elderly woman who has the ability to ascend and descend the stairs in her home will have no difficulty going shopping or attending social activities. Whereas, a frail elderly woman who lost her ability to ascend and descend the stairs in her home may no longer be able to go shopping or attend social activities in buildings without elevators, which may eventually lead to isolation and depression (Rigby & Letts, 2003). Moreover, Lawton discussed that as people age, their competence tends to decrease, and thus, their interaction with the environmental demands becomes even less adaptive and their affect becomes more negative.

According to the association explained on by Lawton (1982; 1983; 1986), when considering older adults with major depression, we come to realize that this association can be much more significant. That is, not only may older adults with major depression lose their competence and adapt less to the environmental demands as they age, but they also may exhibit several symptoms that are inherently debilitating without even considering any of the additional complications imposed by the environment or the aging process. First, according to the DSM-IV-

TR (American Psychiatric Association, 2000), older adults with major depression exhibit affective symptoms such as depressed mood, feelings of guilt and hopelessness, apathy, and loss of interest, pleasure and motivation. Second, older adults with major depression exhibit somatic symptoms, such as psychomotor retardation, fatigue, general weakness, gain or loss in body weight, restlessness, and sleep disturbances (American Psychiatric Association, 2000) that can significantly reduce their physical capacity to perform usual daily activities. Third, older adults with major depression often experience several cognitive symptoms, such as difficulty in thinking, problem-solving, decision making, concentration, and memory (American Psychiatric Association, 2000). Not only can these symptoms result in low self-esteem and diminished physical and cognitive energy (Armenian, Pratt, Gallo & Eaton, 1998; Bruce, 2000; Fried & Guralnik, 1997; Gurland, Wilder, & Berkman, 1988; Penninx et al., 1998), but they may also make coping with additional demands, such as those imposed by the environment, even more challenging. Finally, given that major depression is unpredictable and episodic in nature, living with these symptoms can become even more debilitating (American Psychiatric Association, 2000). Hence, the interaction among competence, aging, environmental demands, and symptoms of depression is complicated. However, the critical association between competence and environmental demands in older adults with major depression has never been explored.

Major depression and the environment are two important factors to be investigated in the rapidly growing and healthcare-demanding older adult population. Older women seem to be most affected by major depression and its disabling effect on ADL and IADL given the responsibility of the daily role-related tasks (Brines, 1994; Marini & Shelton, 1993; Shelton & John, 1996). Understanding the influence of the environment on daily task performance may help in modifying certain features in the environment that can facilitate maximal independence

in FM, ADL and IADL (Shumway-Cook et al., 2002; Yerxa & Baum, 1987). Although symptoms of major depression and age-related impairments can be remediated to a certain extent, modifying the environment may restore competence and intervention may not provide additional benefit. One way to understand the influence of the environment and time on task performance is to examine task performance in two different environments as well as over time. Therefore, in this study we reviewed the literature to (a) locate and evaluate evidence on the influence of the clinic and the home environment on task performance in older adults with major depression, and (b) locate and evaluate evidence of the effect of major depression on task performance over time. The influence of each factor will be reviewed separately.

2.2 METHODS

2.2.1 The influence of the environment on task performance

The databases MEDLINE, PubMed, PsycINFO, and CINAHL were used to search for studies that examined the influence of different environments, namely clinic and home, on task performance in older adults. We used the keywords: *limitations*, *functional limitations*, *task performance*, *performance assessment*, *depression*, and *environment* in combination with *activities of daily living*, *instrumental activities of daily living*, *clinic*, *home*, *older adults*, and *older women*. Additional studies were identified by personal communication with local experts (M. B. Holm, May, 2007 & K. Raina, May, 2007) in functional assessment and by reviewing the citations in the retrieved studies.

2.2.1.1 Relevant Studies

Our search yielded 10 studies (Andrews & Stewart, 1979; Benjamin, 1976; Darragh, Sample, & Fisher, 1998; Haworth & Hollings, 1979; Nygård, Bernspång, Fisher, & Winblad, 1994; Park, Fisher, & Velozo, 1994; Raina, Rogers, & Holm, 2007; Rogers et al., 2003; Sheikh et al., 1979; West et al., 1997). All 10 studies compared task performance of older adults with different diagnoses in the clinic and the home. Also, a literature review was found (Bottari, Dutil, Dassa, & Rainville, 2006) that investigated the evidence provided by other studies regarding the influence of the environment on task performance. Three studies (Darragh et al., 1998; Nygård et al., 1994; Park et al., 1994) utilized the Assessment of Motor and Process Skills [AMPS] (Fisher, 1995) as the outcome measure. These studies were excluded from our review because the AMPS measures task-related impairments in motor and process skills rather than task limitations (e.g., task independence vs. dependence). The literature review by Bottari et al. was also excluded as it included studies that used the AMPS.

To facilitate organization of search findings, we made two strategic decisions. First, we decided to consistently use the terms task performance and task difficulty to refer to the execution of a task and task dysfunctions, respectively. It became apparent from our search that many different terms were used to indicate "the functional ability to perform ADL and IADL" and sometimes multiple terms were used interchangeably in the same study. Terms such as ADL functioning (Andrews & Stewart, 1979); functional level, competence, independence (Haworth & Hollings, 1979); ability to perform (Sheikh et al. 1979); functional status, ability (West et al., 1997), and functional capacity (Benjamin, 1976) were used. Thus, we chose to use the term *task performance* (Raina et al., 2007; Rogers et al., 2003). We will use the term *task performance* to

indicate the positive aspect of task execution (i.e., successful performance), and the term *task difficulty* to indicate the negative aspect of task execution (i.e., inadequate performance).

Second, we will identify the conclusions about the influence of the environment on task performance at three levels: 1) the overall level, that is the global/general evaluation of all tasks grouped together (total score); 2) the domain level, that is, the evaluation of a group of tasks that are similar in nature, such as the domain of FM (domain score), and 3) the task level, that is, the evaluation of individual tasks (task score). This scheme facilitates synthesizing study results because studies compared task performance at different levels and conclusions could differ depending on the level examined.

In the following sections we will summarize the seven studies by level of comparison. Afterwards, the results of the studies will be discussed and our conclusion and recommendation will be provided.

2.2.1.2 Performance Compared at the Overall level (One Study)

Sheikh et al. (1979) used a sample of patients with stroke (n = 73) [age and gender were not provided] to assess independence in task performance in 17 tasks [8 ADL, 3 IADL, 6 FM] in clinic and home (see Table 2-1). Performance-based assessments were conducted at both settings to assess independence in the tasks. The independence scores were compared between the two settings at the overall level. The results showed that patients' total scores indicated that performance was less difficult in the clinic than the home (i.e., clinic was more facilitative than home). Although the same scale was used in the clinic and home, and the same tasks were assessed in both settings, the tasks were not operationally defined for either setting. Thus, it was not clear if exactly the same tasks were assessed in both settings.

2.2.1.3 Performance Compared at the Domain Level (One Study)

Rogers et al. (2003) utilized the standardized performance-based measure Performance Assessment of Self-Care Skills [PASS] (Rogers & Holm, 1989) to assess the task performance of 57 community-dwelling older women with knee osteoarthritis (mean age = 81 years) in the clinic and home, and then determined the percent of agreement at the domain level. The PASS consists of 26 tasks [3 ADL, 18 IADL, 5 FM] (see Table 2-1) categorized into four domains: FM, ADL, cognitive-oriented IADL [CIADL], and physical-oriented IADL [PIADL]). The agreement between clinic and home assessments at the domain level was determined by the number of items where both assessments agreed divided by the total number of items in that domain (e.g., if the agreement between the assessment in clinic and the assessment at home for the FM domain was 3 of 5 items; there was a 60% agreement). Rogers et al. showed that at the domain level, the highest agreement between clinic and home scores was in the FM domain (73.3%) followed by ADL (57.3%), PIADL (55.4%), and CIADL (52%) indicating that clinic and home assessments matched (i.e., were equally demanding) only about 50% of the time in three of the four domains. When the two assessments did not agree, the home was generally less demanding than the clinic.

2.2.1.4 Performance Compared at the Task Level (Four Studies)

Benjamin (1976), with a sample of patients with stroke (the number of patients, gender and age were not provided) assessed independence in 16 tasks [6 ADL, 4 IADL, 6 FM tasks] in the clinic and home (see Table 2-1). Task independence was assessed with the standardized Northwick Park Index in both settings and the scores were compared at the task level. Benjamin used the same tasks, rating scale, and assessment method (performance-based assessment) in the clinic and home. The results showed that differences between the clinic and home scores were small and not statistically significant.

Andrews and Stewart (1979) used a sample of 29 males and females with stroke (age was not provided) to assess task performance of eight tasks [6 ADL, 2 FM tasks] (see Table 2-1) in the clinic and home. A rating scale was used to collect data about the level of assistance needed in the performance of the eight tasks and the scores of the two assessments were compared at the task level. The results showed that more than 50% of the patients performed all tasks (except bathing) with equal difficulty in the clinic and home (see Table 2-1).

Although Andrews and Stewart used the same tasks and the same rating scale in both the clinic and home, they used different assessment methods-- performance-based assessment in the clinic and proxy-report at home. Moreover, Andrews and Stewart named but did not operationally define the tasks, and thus, it was not clear if the tasks were the same in both settings.

Haworth and Hollings (1979), with a sample (n = 38) with rheumatoid arthritis (males = 8, females = 30; mean age = 54.7 years) assessed task performance in 36 tasks [6 ADL, 13 IADL, 9 FM, 5 social/leisure, 3 communication] in the clinic and home (see Table 2-1). A rating scale was used to collect data on the level of assistance needed while performing the tasks and the scores were then compared between the clinic and home at the task level. The same tasks, rating scale, and assessment method (performance-based assessment) were used in the clinic and home. The results showed that in 29 of the 36 tasks assessed the proportion of patients whose performance did not differ between settings exceeded 80%. However, it should be noted that Haworth and Hollings compared the performance of the tasks *visit friends, go out to public places,* and *transfer in/out car seat* in the clinic and the home although the items do not necessarily take place in either setting. The authors failed to comment on how these tasks were assessed; thus calling into question the assessment methodology and the findings.

West et al. (1997) used a sample (n = 97) with visual impairments, ranging in age from 65 to 84 years (males = 36, females = 61; mean age was not provided), to assess task performance in six tasks [4 IADL, 2 FM] in the clinic and home [performance-based assessment was used in each setting] (see Table 2-1). Six standardized clinical tests were used to assess patients' performance and the two assessments were compared at the task level. The results showed that for all six tasks the correlations between task performance in the clinic and home were moderate to high [Pearson correlation coefficient (*r*) ranging from 0.52 to 0.86] indicating similar performances in the two settings.

2.2.1.5 Performance Compared at the Task, Domain, and Overall levels (One Study)

Raina et al. (2007), utilized the PASS to assess independence in task performance for 55 older women with heart failure (mean age = 78.3 years) in the clinic and home and then compared task performance at the task, domain, and overall levels. Raina et al. showed that independence scores did not differ significantly between the clinic and home at the overall level. However, at the domain level independence scores were not significantly different for PIADL, were significantly better at home for CIADL, and were significantly better in the clinic for FM and ADL. Furthermore, at the task level, independence scores for stair use (FM) and trimming toenails (ADL) tasks were significantly better in the clinic. For telephone use, small repairs, home safety (CIADL) and cleanup after meal preparation (PIADL) tasks were significantly better at home; and on the rest of the PASS tasks the scores were not significantly different between clinic and home. Hence, the environment had no effect on overall task performance but did impact performance at the domain and task levels.

2.2.2 The effect of major depression on task performance over time

We used MEDLINE, PubMed, PsycINFO, CINAHL, and Cochrane databases to locate studies that examined major depression and disability (difficulties in ADL and IADL), over time, in older adults aged 70 years or older. Studies had to have included samples of older adults and be written in English. Additional studies were identified by Margo B. Holm (personal communication, May, 2007), hand-searching issues of relevant journals; and by reviewing the citations in the retrieved studies. The search keywords were: *disability*, *functional disability*, *physical disability*, *longitudinal*, *difficulties in task performance*, *depression*, *late life depression*, and *depressive symptoms*, in combination with *activities of daily living*, *instrumental activities of daily living*, *older adults*, and *older women*. In the following three sections we report the results of our search and discuss in detail studies that are most relevant to our aim *(Relevant Studies*), discuss these results and identify gaps found in the literature *(Discussion)*, and conclude our study *(Summary)*.

2.2.2.1 Relevant Studies

Numerous studies (n = 39) have examined major depression and disability in older adults. The relevance of these studies for the current study, which seeks to understand the performance of specific daily living tasks in the presence of major depression in older women, is limited because:

- They used a cross sectional design (Laukkanen, Heikkinen, Schroll, & Kauppinen, 1997; Mulsant, Ganguli, & Seaberg, 1997; Seaburn, Lyness, Eberly, & King, 2005; Steffens, Hays, & Krishnan, 1999), and hence did not provide information on task performance over time;

- They examined global disability (Beekman et al., 1997; Hays Wells, Sherbourne, Rogers, & Spritzer, 1995; Heikkinen, Berg, Avlund, & Törmäkangas, 2002; Wells, Burnam, Rogers, Hays, & Camp, 1992), and hence did not provide information about specific daily living activities;
- They measured physical impairments such as balance and strength (Callahan et al., 1998; Everson-Rose et al., 2005; Schroll, Avlund, & Davidsen, 1997), and hence did not provide information about task performance;
- They studied samples where depression was secondary to a physical diagnosis such as stroke (Armenian, Pratt, Gallo, & Eatson, 1998; Dunlop, Manheim, Song, Lyons, Rowland, & Chang, 2003; Forsell, Jorm, & Winblad, 1998; Gill, Allore, Holford, & Guo, 2004; Greelings, Beekman, Deeg, Twisk, & van Tilburg, 2001; Guccione et al., 1994; Han, 2002; Han & Jylha, 2006; Hays et al., 1995; Hays, Saunders, Flint, Kaplan, & Blazer, 1997; Kiosses & Alexopoulos, 2005; Kurland, Gill, Patrick, Larson, & Phelan, 2006; Lenze et al., 2005; Ostir, Volpato, Kasper, Ferrucci, & Guralnik, 2001; Turner & Noh, 1988; Vaccarino, Kasl, Abramson, & Krumholz, 2001; Whyte, Mulsant, Vanderbilt, Dodge, & Ganguli, 2004) and hence, did not provide information on disability in the presence of a primary diagnosis of major depression;
- They used screening tests (Bruce, Seeman, Merrill, & Blazer, 1994; Gill et al., 2004; Guccione et al., 1994; Han, 2002; Han & Jylha, 2006; Hays et al., 1997; Kurland et al., 2006; Lenze et al., 2001; Lenze et al., 2005; Ostir et al., 2001; Turner & Noh, 1988; Whyte et al., 2004) as opposed to diagnostic tests for identifying depression, and hence, although symptoms of depression were present, a definitive diagnosis of major depression was not made;

- They were epidemiologic studies (Bowling, Farquhar, & Grundy, 1996; Chou, 2005; Gallo, Rabins, Lyketsos, Tien, & Anthony, 1997; Gallo, et al., 2003; Heikkinen & Kauppinen, 2004; Kennedy, Kelman, & Thomas, 1990; Laukkanen et al., 1994; Penninx, Leveille, Ferrucci, van Eijk, & Guralnik, 1998; Prince, Harwood, Thomas, & Mann, 1998; Williamson & Schulz, 1992; Zeiss, Lewinsohn, Rohde, & Seeley, 1996) which examined the emergence of depression and coincident disability in populations, and hence, did not chart change in disability.

Consequently, there is a dearth of information about the specific tasks that are most affected by major depression and the course of task performance as symptoms are alleviated and mood stabilizes.

Only two studies provide data on the trajectory of specific tasks, and ADL/IADL domains during recovery from depression (Table 2-2). Over 6 months, Rogers, Holm, Goldstein, McCue, and Nussbaum (1994) found that patients self-reported their performance as stable, while performance testing of task abilities detected declines in domain scores related to FM, ADL, and IADL. Over 1 year, Hays, Steffens, Flint, Bosworth, and George (2001) found that some tasks remained stable, some improved, and some got worse.

The first of the two relevant studies, Rogers et al. (1994), assessed the task performance of 9 subjects with major depression 2 weeks and 6 months after discharge. Depression was identified with the diagnostic criteria of the Schedule for Affective Disorder and Schizophrenia-Lifetime (SADLs-L) interview which was conducted by a qualified clinician. Difficulties in functional performance were assessed with a) self-report using the Activities of Daily Living Scale from the Older Americans Resources and Services Multidimensional Functional Assessment (OARS-ADL), b) informant or proxy-report also using the OARS-ADL, and c)

version 1.0 of the Performance Assessment of Self-Care Skills [PASS] (Rogers, 1987). The OARS-ADL scale contains task of both ADL and IADL, with each task rated on a 3-point scale (0 = significant disability, and 3 = independent performance). The PASS is a standardized performance-based assessment tool that assesses performance in 25 ADL/IADL categorized into three functional domains. The PASS domains are (Functional Mobility (FM): move from prone to supine position and rise from bed, sit and rise from a chair, lift 3-lb object from floor, ascend and descend stairs, enter a doorway by using a key, locate eating, sleeping, and toileting areas; Personal Care Activities of Daily Living (PC): feed self, select appropriate clothing, don clothing, brush teeth, groom hair, groom fingernails and toenails, bathe self; and IADL: wash dishes, make a bed, wash clothing, clean a floor with an electronic broom, verbalize appropriate response to danger, use telephone to obtain information, sew on a button, simulate shopping by selecting and purchasing gloves, balance a checkbook after writing a check, prepare an envelope for mailing the check, cook pudding on a range, demonstrate management of medication). Although 25 tasks were assessed, only the **domain scores** were compared between the two assessments

The results showed that both the patients (self-report) and informants (proxy-reports) reported stable performance of ADL and IADL tasks over time. However, that PASS domain scores declined over time in FM and IADL but not in PC.

The second relevant study, Hays et al. (2001) tested the hypothesis that for older adults with depression initial severity of depressive symptoms (i.e., at baseline) predicts functional decline (i.e., difficulties in ADL/IADL) 1 year later (i.e., follow-up), even after controlling for demographic factors, clinical features of initial episode, and improvement in depressive symptoms during follow-up. Hays et al. used a mixed sample of 113 outpatients and inpatients

(females = 69%, males = 31%; mean age = 69.5 years); all had clinically significant depressive symptoms or a previous diagnosis of mood disorder. Additionally, 98% of the subjects were treated with naturalistic treatment (not described) determined by patients' clinical status rather than by a fixed protocol. All subjects were screened with the CES-D for depressive symptoms to determine their eligibility to participate in the study. Only subjects with a score of 16 (or above) or with major depression diagnosis were eligible. At baseline, the Duke Depression Evaluation Schedule [a composite diagnostic instrument that includes the depression assessment section of the National Institute of Mental Health (NIMH) Diagnostic Interview Schedule] was administered. Difficulties in ADL/IADL (at follow-up) were identified with self-reports on 7 ADL (eating, dressing/undressing, grooming, walking, bathing/showering, using the toilet, bending down) and 9 IADL (walking 0.25 mile, walking up/down stairs, getting around the neighborhood, shopping for groceries, keeping track of money/bills, taking care of children, cleaning house, preparing meals, doing yard work/gardening). Difficulties in ADL/IADL were compared between baseline and follow up at the **domain**, and **task levels.**

The results at follow-up showed that at the **domain level,** 10.6% of the sample (n = 113) had increased disability in ADL, with the greatest increases seen in walking, bending down, and dressing/undressing. Similarly, 23% had increased disability in IADL specifically in walking ¼ mile, walking up and down stairs, keeping track of money, and cleaning house. However, 15.1% had less ADL disability with greatest improvements seen in walking and bending down. In addition, 46% had less IADL disability at 1 year with greatest improvement seen in taking care of children and preparing meals. Nonetheless, despite these declines and improvements, the greatest proportion of the sample remained stable over 1 year in ADL (74.3%), with a smaller proportion stable in IADL (31%). Finally, at the **task level** the results showed that the proportion

of patients who exhibited less disability in basic ADL after 1 year were not significantly higher than those who exhibited more disability. However, for IADL, the proportion of patients who exhibited less disability in IADL after 1 year was significantly greater than those who exhibited more disability. The only exceptions were walking ¼ mile and walking up/down stairs. The percentage of patients who exhibited **more disability** after 1 year in basic ADL and IADL were as follow: eating (1.8%), dressing/undressing (4.4%), grooming (3.5%), walking (7.1%), bathing/showering (1.8%), using the toilet (2.7%), bending down (5.3%), walking ¼ mile (8.8%), walking up/down stairs (8%), getting around the neighborhood (5.3%), shopping for groceries (4.4%), keeping track of money/bills (7.1%), taking care of children (6.2%), cleaning the house (7.1%), preparing meals (4.4%), and doing yard work/gardening (5.3%). The percentage of patients who exhibited **less disability** after 1 year in ADL and IADL and they were as follows (significant for all tasks except walking .025 mile and walking up/down stairs): eating (2.7%), dressing/undressing (6.2%), grooming (5.3%), walking (8.8%), bathing/showering (6.2%), using the toilet (7.1%), bending down (12.4%), walking ¼ mile (10.6%), walking up/down stairs (14.2%), getting around the neighborhood (13.3%), shopping for groceries (14.2%), keeping track of money/bills (16.8%), taking care of children (21.2%), cleaning the house (18.6%), preparing meals (20.4%), and doing yard work/gardening (16.8%). The percentage of patients who showed **stable performance** over time were as follows: eating (95.6%), dressing/undressing (89.4%), grooming (91.2%), walking (84.1%), bathing/showering (92%), using the toilet (90.3%), bending down (82.3%), walking ¼ mile (80.5%), walking up/down stairs (77.9%), getting around the neighborhood (81.4%), shopping for groceries (81.4%), keeping track of money/bills (76.1%), taking care of children (72.5%), cleaning the house (74.3%), preparing meals (75.2%), and doing yard work/gardening (77.9%).

2.3 DISCUSSION

2.3.1 The influence of the environment

A large number of older adults, especially older women, experience limitations in task performance. This number will grow significantly in the future, given the changing demographics. Investigating the factors that may contribute or lead to these limitations and intervening to remediate them may aid in preventing or minimizing the limitations. In this study, we searched the literature for studies that investigated the influence of the environment on task performance as a contributing factor to limitations in older adult populations. We found 7 relevant studies. The studies provided conflicting results regarding the influence of environment on task performance of older adults. At the overall level, one study (Sheikh et al. 1979) found that home was more demanding than clinic while another study (Raina et al., 2007) showed that clinic and home were equally demanding (see Table 2-1). At the domain level, one study (Rogers et al., 2003) showed that clinic and home were equally demanding most of the time while another study (Raina et al., 2007) showed that for the CIADL domain clinic was more demanding than home and for FM and ADL domains, home was more demanding than clinic (see Table 2-1). Finally, At the task level, four studies (Andrews & Stewart, 1979; Benjamin, 1976; Haworth & Hollings, 1979; West et al., 1997) showed that clinic and home were equally demanding, and one study (Raina et al., 2007) showed that for telephone use, home safety, small repairs (CIADL), and cleanup up after meal preparation (PIADL) clinic was more demanding than home, while for stair use and trimming toenails home was more demanding than clinic (see Table 2-1). In summary, there are conflicting results at the overall, domain, and task levels regarding the equivalence of environmental demands between the clinic and home.

We had four major concerns regarding the evidence derived from the available literature on the influence of environment on task performance in older adults. First, although three of the seven studies (Benjamin, 1976; Raina et al., 2007; Rogers et al., 2003) used standardized assessment tools, only two (Raina et al., 2007; Rogers et al. 2003) used a standardized assessment tool (PASS; Rogers & Holm, 1989) that was reliable and valid, comprehensively assessed task performance in different functional domains, and most importantly, allowed for the comparison of task performance at three distinct levels (overall, domain, and task levels). Second, 50% of the studies had methodological limitations such as differences in data collection methods between settings (Andrews & Stewart, 1979), poor definition of tasks assessed (Andrews & Stewart, 1979; Haworth & Hollings, 1979; Sheikh et al. 1979), or lack of demographic information (Andrews & Stewart, 1979; Benjamin, 1976; Sheikh et al., 1979). Third, all seven studies used samples with physical disabilities such as stroke (Andrews & Stewart, 1979; Benjamin, 1976; Sheikh et al., 1979), heart failure (Raina et al., 2007), knee osteoarthritis (Rogers et al., 2003), rheumatoid arthritis (Haworth & Hollings, 1979), and visual impairments (West et al., 1997). Hence, no study used a sample of patients with a seriously debilitating mental illness such as major depression. Fourth, all studies except two (Raina, Rogers, & Holm, 2007; Rogers et al., 2003) used mixed samples of males and females and none conducted a gender analysis to determine gender differences. In addition, in the two studies that used samples of women only (Raina, et al., 2007; Rogers et al., 2003), the primary diagnosis of the samples was not major depression

2.3.2 The influence of major depression over time

Only two studies were found that were directly relevant to our aim of finding evidence on the effect of major depression over time on task performance (Hays et al., 2001; Rogers et al., 1994), see Table 2-2. The two studies were similar in some aspects and different in others. Both studies used samples of older adults with major depression with no comorbid medical illness, used measures of difficulties in ADL and IADL, and specified which tasks (Hays et al.) and domains (Hays et al.; Rogers et al.) of ADL and IADL were affected by the course of depression. However, the samples used in these two studies included both males and females. Moreover, while Hays et al. used only self-report to assess task performance in ADL and IADL, Rogers et al. used three assessment methods to assess task performance in ADL and IADL (self-report, informant, and performance-based assessment). The results of the common assessment method (self-report) between the two studies showed a stable performance over time for ADL and stable (Hays et al. & Rogers et al.), declined, or improved performance for IADL (Hays et al). Finally, while Rogers et al. assessed task performance at the domain level; Hays et al. assessed task performance at both the domain and task level.

Given the aforementioned gaps identified in the literature related to major depression and disability, and the mixed results of the study by Hays et al. (2001), our aim was to examine changes in ADL and IADL, over time, in older women diagnosed with major depression. We are not putting forth hypotheses because this study is exploratory as the research on samples with physical impairments may not translate well to those with affective impairments. While the study by Rogers et al. (1994) would lead us to conclude that our subjects will decline, subjects in the Rogers et al. study were closer to an acute depressive episode and hospitalization. Similarly, while the Hays et al. study would lead us to conclude that the majority of our subjects would be

stable in ADL, and a smaller proportion would decline, improve, or remain stable in the performance of IADL, subjects in the Hays et al. study self-reported performance, whereas those in our study will be performance-tested. There are known difference in results attributable to assessment methods (Rogers et al., 2003; Rogers, Holm, & Stone, 1997).

2.4 SUMMARY

A large number of older adults with major depression, specifically women, exhibit significant limitations in ADL and IADL, and these limitations seem to increase over time. Yet no studies have investigated the influence of environment, or major depression over time, on the functional performance of daily activities in older adults with major depression. Moreover, although women are twice as vulnerable to major depression as men, no studies have investigated the influence of environment or time on task performance in women with major depression. Hence, the purpose of the study in Chapter 3 was to investigate the influence of environment (clinic versus home) on the performance of ADL and IADL in older women with major depression. Likewise, the purpose of the study in Chapter 4 was to investigate the influence of time on the performance of ADL and IADL in older women with major depression.

Table 2-1 Summary of Findings on Task Performance in the Clinic and Home

Study	Clinic > Home	Home > Clinic	Clinic = Home	Methods	Task [Domain]	Authors' Conclusion
Benjamin (1976)			All tasks--- No tasks were specified to be better at either place	Clinic: Performance-based assessment (Northwick Park Index) Home: The same Performance-based observation used in clinic	Tasks: 16 activities [6 ADL, 4 IADL, 6 FM] ADL: Dressing, Bathing/washing, Toileting, Grooming, Feeding, Continence IADL: Preparation of tea-making, Making tea, Cooking, Use of taps FM: Transfer (bed/chair), Transfer (floor/chair), Mobility indoors, Mobility outdoors, Stairs (up), Stairs (down)	"The differences which occurred throughout the evaluation procedures were small and proved to be statistically insignificant"p.303
Sheikh et al. (1979)	Overall clinic scores were better than the overall home scores— differences are in activities involving equipment (such as cookers) – no other tasks were specified because comparison was done at the overall level.			Clinic: Performance-based observation Home: The same Performance-based observation used in clinic	Tasks: 17 tasks [8 ADL, 3 IADL, 6 FM] ADL: Dressing, Washing, Bathing, Using lavatory, Continence, Grooming, Brushing teeth, Feeding IADL: Preparing for making tea, Making tea, Using taps FM: Transfer from floor to chair, Transfer from chair to bed, Walking indoors, Walking outdoors, Ascending a flight of stairs, Descending a flight of stairs Overall: Home total scores were higher (worse) than clinic total scores	"There is a slight tendency for home scores to be a little higher [higher meaning worse performance] than [clinic] scores" ($p<.05$). p. 54
Andrews and Stewart (1979)	Getting into bed (34%) Wash hands and face (41%) Bathing (45%) Dressing (41%) Putting on shoes (34%) Feeding (41%) Toileting (31%) Getting around inside house (24%)	Getting into bed (0%) Wash hands and face (3%) Bathing (14%) Dressing (3%) Putting on shoes (0%) Feeding (7%) Toileting (7%) Getting around inside house (7%)	Getting into bed (66%) Wash hands and face (56%) Bathing (41%) Dressing (56%) Putting on shoes (66%) Feeding (52%) Toileting (62%) Getting around inside house (69%)	Clinic: Performance-based observation Home: Proxy-report	Tasks: 8 tasks [6 ADL, 2 FM] ADL: Wash hands and face, Bathing, Dressing, Putting on shoes, Feeding, Toileting, FM: Getting into bed, Getting around inside house	"Over one third of the patients did better in the clinic in all activities"p.45

Note. % = percentage of patients; ADL s= activities of daily living; IADL = instrumental activities of daily living.

Study	Clinic > Home	Home > Clinic	Clinic = Home	Methods	Task [Domain]	Authors' Conclusion
Haworth and Hollings (1979)	<u>In terms of tasks:</u> In 16% of the tasks (6 of 36) patients were significantly better in the clinic. These tasks were: Indoor tasks: Do laundry: 7 patients ($p<.05$) Housework: 9 patients ($p<.05$) Bath/shower: 8 patients ($p<.01$) Outdoor tasks: Transfer in/out car: 11 patients ($p<.01$) Shopping: 12 patients ($p<.01$) Go out to public places: 14 patients ($p<.001$). <u>In terms of patients</u> 68% of the patients did better in the clinic (i.e., performance deteriorated at home)	<u>In terms of tasks:</u> In 10% of the tasks (4 of 36) patients were significantly better at home. These tasks were: Indoor tasks: Do laundry: 0 patients ($p<.05$) Housework 1 patient ($p<.05$) Bath/shower 0 patients ($p<.01$) Outdoor tasks: Transfer in/out car :1 patient ($p<.01$) Shopping: 1 pt ($p<.01$) Go out to public places: 1 pt ($p<.001$). <u>In terms of patients</u> 10% of the patients did better (show improvement) at home	<u>In terms of tasks:</u> In 29 of 36 tasks (80%) the proportion of patients who scored the same in both settings exceeded 80%) <u>In terms of patients:</u> 22% of the patients did the same in clinic and home	<u>Clinic</u>: Performance-based assessment <u>Home</u>: The same Performance-based observation used in clinic	<u>Tasks:</u> 36 tasks: [6 ADL, 13 IADL, 9 FM, 5 social/leisure, 3 communication], <u>ADL</u>: Use toilet, Wash upper body, Wash all over body, Dress, Use bath/shower, Eat and drink <u>IADL</u>: Pick up articles from floor, Manage money, Reach forwards, Reach above shoulders, Sew, Do house repairs, Wash up dishes, Turn taps/keys/switches, Do laundry, Prepare all meals, Cook all meals, Do housework, Carry shopping <u>FM</u>: Use wheelchair inside, Use wheelchair outside, Walking inside house, Transfer up/down from chair, Use public transport, Drive car, Go up/down stairs, Walk outside house, Transfer in/out car seat <u>Social Activities</u>: Turn on/off TV/radio, Entertain, Pursue interest/hobby, Visit friends/relatives, Go out to public places <u>Communication</u>: Read book, Use telephone, Write	"In our prospective study of patients with rheumatoid arthritis a predischarge ADL assessment has been compared with an identical assessment at home ten days later. This revealed that most activities tended to be carried out with comparable facility…" p. 61

Note. ADL= activities of daily living; IADL = instrumental activities of daily living.

Study	Clinic > Home	Home > Clinic	Clinic = Home	Methods	Task [Domain]	Author's conclusion
West et al., (1997)	Stair climb: 7% of the patients (r=.77)* Stair descend: 12% of the patients 41, (r=.76)* Plug insertion: 33% of the patients. Moreover, patients who had poor lighting at home (27%) did better in the clinic Look up a number: 45% of the patients did better in the clinic. Moreover, those who took less than 25 seconds to read at home (% not specified) did better in the clinic Dial a telephone number: 44% of the patients did better in the clinic. Reading test: Patients without visual impairments: 91% of the patients did better in the clinic	Stair climb: 93% of patients (r=.77) Stair descend: 88% of the patients (r=.76) Plug insertion: 77% of the patients (r=.56) Look up a telephone number: 55% of the patients did better at home. moreover, patients who were unable to perform the task in the clinic and those who took longer than 25 seconds in the clinic, did better at home (% not specified), those who were visually impaired did better at home (62%) –not significant. Patients who used magnifiers at home did better at home (r=.62) Dial a telephone number: 56% of the patients did better at home. Moreover, those who took longer than 7 seconds (% not specified) and those with visual impairments (60%) did better at home—not significant (r=.60) Reading test: Patients without visual impairments: 9% of the patients did better at home	Reading test: Patients with visual impairments 50% did better in the clinic and 50% did better at home. (r=.75 for slow reader, r=.86 for fast reader)	Clinic: Performance-based observation Home: The same Performance-based observation used in clinic	Tasks: 6 tasks [4 IADL, 2 FM] IADL: Plug insertion, Reading, Look up a telephone number, Dial a telephone number, FM: Stair climb, Stair descend	"There was good correlation between the scores on selected tasks performed at home and at the clinic, the correlation coefficients ranged from .52 to .86. Where performance was consistently better at home such as on the task of stair climbing, the ranking of participants on performance at either side was similar. It appears that a familiar environment and use of usual adaptive mechanisms may have led to an improvement of performance at home in those who generally performed slowly" p. M214

Note. r = Pearson correlation coefficients indicating the association between performance in the clinic and performance at home; * = patients who had stairs at home did better in clinic than those who did not have stairs at home; ADL = activities of daily living; IADL = instrumental activities of daily living.

Study	Clinic > Home	Home > Clinic	Clinic = Home	Methods	Task [Domain]	Author's conclusion
Rogers et al. (2003)	 <u>Domain level:</u> FM: 5.6% ADL: 10.5% CIADL: 5.9% PIADL: 7.5%	 <u>Domain level:</u> FM: 20.7% ADL: 32.2% CIADL: 38.1% PIADL: 35.1%	<u>Overall level:</u> Significant main effect for methods (ANOVA p value) <u>Domain level:</u> FM: 73.3% PBO-clinic = PBO-home ($p<.001$) ADL: 57.3% PBO-clinic = PBO-home ($p<.05$) CIADL:55.4% PBO-clinic = PBO-home ($p<.001$) PIADL 52% PBO-clinic = PBO-home ($p<.05$)	<u>Clinic:</u> Performance Assessment of Self-Care Skills (PASS) <u>Home:</u> PASS	<u>Tasks:</u> 26 tasks: [3 ADL, 14 CIADL, 4 PIADL, 5 FM] <u>ADL</u>: Oral hygiene, Trimming toenails, Dressing <u>CIADL</u>: Shopping, Bill paying by check, Checkbook balancing, Mailing bills, Telephone use, Medication management, Obtaining critical information-auditory, Obtaining critical information-visual, Small repairs, Home safety, Oven use, Stovetop use, Use of sharp utensils, Playing bingo <u>PIADL:</u> Cleanup after meal preparation, Sweeping, Carrying the garbage, Changing bed linens <u>Domain:</u> Functional Mobility (FM), Personal Care Activities of Daily Living (ADL), Cognitive-Instrumental Activities of Daily Living (CIADL), Physical-Instrumental Activities of Daily Living (PIADL)	"Limitations estimates based on self-reports, proxy-reports, clinical judgments, and hospital performance-based assessments are not interchangeable with in-home task performance" p. 640
Raina, Rogers and Holm (2007)	<u>Task level:</u> Stair use (FM) Trimming toenails (ADL) <u>Domain level:</u> FM and ADL <u>Overall level:</u> Not applicable	<u>Task level</u> Telephone, Home safety, Small repairs (CIADL), Cleanup after meal preparation (PIADL) <u>Domain level:</u> CIADL <u>Overall level:</u> Not applicable	<u>Task level:</u> All other tasks <u>Domain level:</u> PIADL <u>Overall level:</u> scores are the same at the overall level	<u>Clinic:</u> PASS <u>Home:</u> PASS	<u>Tasks:</u> 26 tasks: [3 ADL, 14 CIADL, 4 PIADL, 5 FM] <u>ADL</u>: Oral hygiene, Trimming toenails, Dressing <u>CIADL</u>: Shopping, Bill paying by check, Checkbook balancing, Mailing bills, Telephone use, Medication management, Obtaining critical information-auditory, Obtaining critical information-visual, Small repairs, Home safety, Oven use, Stovetop use, Use of sharp utensils, Playing bingo <u>PIADL:</u> Cleanup after meal preparation, Sweeping, Carrying the garbage, Changing bed linens <u>Domain:</u> FM, ADL, CIADL, PIADL	" The influence of environment can be neutral, enabling, disabling depending on the global scores of level of analysis being considered" [therefore] "it would be problematic to make a general statement about the enabling/disabling influence of the environment on the activity performance of community dwelling older women with heart failure"p.19

Note. PBO-clinic= performance-based observation in the clinic; PBO-home = performance-based observation at home; ADL = activities of daily living; IADL = instrumental activities of daily living; CIADL = cognitive-oriented IADL, PIADL = physical-oriented IADL.

Table 2-2 Studies that Examined the Effect of Time on Task Performance in Older Adults with Major Depression

Study	Participants	Study design	Outcomes	Assessment level(s)	Results	Author's conclusion
Rogers, Holm, Goldstein, McCue, and Nussbaum (1994)	9 subjects with major depression Mean age = 74.94 years	Patients were assessed at two points of time (2 weeks and 6 months after discharge)	<u>Depressive symptoms (independent variable):</u> 1. Schedule for Affective Disorder and Schizophrenia-Lifetime (SADS-L) interview 2. Hamilton Depression Scale <u>Functional Performance (dependent variable):</u> 1. ADL/IADL: self-reported using the Activities of Daily Living Scale from Older Americans Resources and Services Multidimensional Functional Assessment (OARS-ADL) 2. ADL/IADL: informant (proxy-report) using the OARS-ADL 3. FM, PC, IADL assessed by an older version of the Performance Assessment of Self-Care Skills (PASS)	<u>Domain level only:</u> Composite score for ADL and IADL in both the OARS and PASS	<u>Functional Performance (dependent variable):</u> 1. <u>ADL/IADL: self-report:</u> Stable performance over time 2. <u>ADL/IADL: informant (proxy-report):</u> Stable performance over time 3. <u>FM, PC, IADL (PASS)</u> Performance significantly worsened over time for FM ($p < .05$) and IADL ($p < .05$). Performance did not significantly worsen on PC over time ($p > .05$).	"Elderly patients with depression may experience subtle deterioration that only becomes apparent on performance tests" p. 914

Note. PC = personal care activities of daily living; IADL = instrumental activities of daily living.

Study	Participants	Study design	Outcomes	Assessment level(s)	Results	Author's conclusion
Hays, Steffens, Flint, Bosworth, and George (2001)	n = 113 subjects with clinically significant major depression Mean age = 69.5 years 69% females 31% males	Prospective cohort design. Participants were followed for 12 months. Clinical assessments were administered at the beginning of the study and at 1-year follow-up.	<u>Depressive symptoms (independent variables)</u>: 1. CES-D screening 2. Duke Depression Evaluation Schedule 3. Hamilton Depression Rating Scale <u>Functional performance (dependent variables)</u>: **7 ADL**: self-reported **9 IADL**: self-reported Change in functional performance after one year was described as : • Decreased over time; • Improved over time; • Stable over time (no difficulty at baseline or after 1 year; OR, difficulty at both baseline and after 1 year).	<u>Task level:</u> ADL = eating, dressing/ undressing, grooming, walking, bathing/ showering, using the toilet, bending down IADL = walking 0.25 mile, walking up/down stairs, getting around the neighborhood, shopping for groceries, keeping track of money/bills, taking care of children, cleaning house, cleaning house, preparing meals, doing yard work/gardening <u>Domain level:</u> ADL: composite score IADL: composite score	**<u>Task level (ADL):</u>** • Decreased functional ability after 1 year (% of patients): Eating 1.8%, dressing/undressing 4.4%, grooming 3.5%, walking 7.1%, bathing/showering 1.8%, using the toilet 2.7%, bending down 5.3% • Improved functional ability after 1 year (% of patients): Eating 2.7%, dressing/undressing 6.2%, grooming 5.3%, walking 8.8%, bathing/ showering 6.2%, using the toilet 7.1%, bending down 12.4% • Stable functional ability after 1 year (% of patients): Eating 95.6%, dressing/undressing 89.4%, grooming 91.2%, walking 84.1%, bathing/showering 92%, using the toilet 90.3%, bending down 82.3% **<u>Task level (IADL):</u>** • Decreased functional ability after 1 year (% of patients): walking 0.25 mile 8.8%, walking up/down stairs 8.0%, getting around the neighborhood 5.3%, shopping for groceries 4.4%, keeping track of money/bills 7.1%, taking care of children 6.2%, cleaning house 7.1%, preparing meals 4.4%, doing yard work/gardening 5.3% • Improved functional ability after 1 year (% of patients): walking 0.25 mile 10.6%, walking up/down stairs 14.2%, getting around the neighborhood 13.3%, shopping for groceries 14.2%, keeping track of money/bills 16.8%, taking care of children 21.2%, cleaning house 18.6%, preparing meals 20.4%, doing yard work/gardening16.8% • Stable functional ability after 1 year (% of patients): walking 0.25 mile 80.5%, walking up/down stairs 77.9%, getting around the neighborhood 81.4%, shopping for groceries 81.4%, keeping track of money/bills 76.1%, taking care of children 72.5%, cleaning house 74.3%, preparing meals 75.2%, doing yard work/gardening 77.9% **<u>Domain level:</u>** ADL: decreased (10.6%), improved (15.1%) stable (74.3%) IADL: decreased (23%), improved (46%%) stable (31%)	"Instrumental support was generally protective against worsening performance on instrumental abilities of daily living among elderly patients with recurrent unipolar depression]" p. 1850

Note. CES-D = center for epidemiologic studies depression scale.

3.0 DIFFERENCES IN TASK PERFORMANCE BETWEEN THE CLINIC AND THE HOME IN OLDER WOMEN WITH MAJOR DEPRESSION

3.1 BACKGROUND AND SIGNIFICANCE OF THE PROBLEM

Major depression is one of the most prevalent and debilitating illnesses worldwide (Murray & Lopez, 1996). Prevalence studies showed that major depression occurs in 6.7% of the United States general population age 18 years and older (Kessler, Chiu, Demler, & Walters, 2005). It affects both men and women and may appear at any age group from childhood to old age (Mirowsky & Ross, 1992). However, it is twice as common in women as in men with an average prevalence of 20% for women and 10% for men (Desai & Jann, 2001). Moreover, it is more prevalent in younger and middle-aged adults [6.6%] than in older adults [4%] (Mojtabai & Olfson, 2004). Persistent sadness, feeling of hopelessness, decreased energy, physical fatigue, and loss of motivation, interest, and pleasure are a few examples of the symptoms of major depression (American Psychiatric Association, 2000). Such symptoms can be disabling for activities of daily living regardless of the age of onset (Guralnik, Leveille, Hirsch, Ferrucci, & Fried, 1997). Although major depression is less prevalent in old age compared to other age groups, the disabling impact of major depression on daily activities is most significant in older adults, given the age-related cognitive and physical decline and the losses in health, marriage, employment, and economic well-being that may be experienced by many older adults (Mirowsky

& Ross, 1992). Furthermore, when coexisting with major depression, age-related chronic illnesses, such as heart conditions, diabetes, and arthritis, can increase the severity of the symptoms of major depression, thereby, accelerating functional decline and leading to greater disability in all health domains (Mirowsky & Ross, 1992). Given that many older adults experience these chronic conditions as well as other social and economic difficulties (Callahan et al., 2005), healthcare professionals often mistakenly conclude that depression is a normal consequence of these difficulties (National Institute of Mental Heath, 2007). As a result, major depression has become a widely underrecognized and undertreated illness in old age (Desai, 2001; Kessler et al., 2003; Sheehan, 2004; Wells et al., 1989).

Other factors that can add to the complications of major depression symptoms in old age include environmental demands, or the challenges imposed by the environment in which daily activities are performed, for example the home. Lawton (1982) was the first to discuss the association between the competence of the individual, that is the individual's ability to perform daily activities, and environmental demands. He discussed that people with high competence can better adapt to environmental demands, whereas people with low competence are less able to adapt to environmental demands. In other words, the greater the ability of the individuals to perform functional mobility (FM), activities of daily living (ADL) and instrumental activities of daily living (IADL), the more independent they are even in a demanding environment. In contrast, the lower the ability of the individuals to perform FM, ADL and IADL, the less independent they are in a demanding environment. For example, a person with good visual abilities will be more independent (will need no assistance) while walking in or outside the house even on a slippery floor. In contrast, a person with poor visual abilities will be less independent (or will need assistance) while walking in or outside the house, on a slippery floor.

Lawton further discussed that this association between competence and adaptation to environmental demands becomes critical in old age, because as people age their competence to perform ADL and IADL tends to decrease, and therefore, their ability to adapt to environmental demands decreases.

Commonly, the environment in which people perform tasks of FM, ADL, and IADL is the home (Yerxa & Baum, 1987). However, when patients in the healthcare system are asked by therapists to perform these tasks as part of a functional assessment, performance often occurs in a standardized environment such as a clinic. Both the home and the clinic may provide certain simplifications or impose certain demands or challenges. Therefore, the performance of FM, ADL, and IADL not only depends on people's competencies and weaknesses, but it also depends on the facilitative or demanding nature of the environment.

Although Lawton discussed the critical association between competence and environment and their relation to old age, this association has not been explored especially in older adults with major depression. This association can be more complicated in people with depression. That is, according to the DSM-IV-TR criteria, people with major depression not only have affective symptoms such as the loss of motivation, apathy, and depressed mood, but they also have somatic symptoms such as fatigue, psychomotor retardation, loss of energy, and restlessness, and cognitive symptoms such as difficulties in decision making, thinking, and concentration (American Psychiatric Association, 2000). These symptoms can be significantly debilitating to the independent performance of daily activities (Armenian, Pratt, Gallo & Eaton, 1998; Bruce, 2000; Fried & Guralnik, 1997; Gurland, Wilder, & Berkman, 1988; Penninx et al., 1998), in that, older adults with major depression may need assistance in initiating, continuing, or completing essential tasks due to these disabling symptoms (Rogers & Holm, 2000). The

unpredictable episodic nature of these symptoms makes the debilitating effects of major depression even more significant in role-related daily activities (American Psychiatric Association, 2000). Furthermore, these symptoms may decrease the person's physical competence and the psychological energy to successfully interact with the surrounding environment, which may eventually lead to functional disability (Simon et al., 1998). In fact, several studies have shown an association between symptoms of major depression and difficulties in daily activities in several older adult populations (Beekman, Deeg, Braam, &Van Tilburg, 1997; Bruce, 1999; Bruce, 2000; Prince, Harwood, Blizard, Thomas, & Mann, 1997; Stuck et al., 1999; Wells et al., 1989). Not only that, but older adults with major depression were found to have difficulties in daily activities significantly more than well older adults or older adults with physical diagnoses (Wells et al., 1989). Finally, since it is important to consider the association between the person's physical and mental health (Berkman et al., 1986; Bruce, 2000), when examining the association between competence and environmental demands, the debilitating nature of the symptoms of major depression, especially in older adults, should be seriously considered.

Therefore, in this study we aim to identify which environment, the clinic or the home, yields greater independence in FM, ADL, and IADL for older women with depression. To achieve our aim we will compare the overall performance and the performance in four functional domains (functional mobility [FM], personal care activities of daily living [PC], cognitively-oriented instrumental activities of daily living [CIADL] and physically-oriented instrumental activities of daily living [PIADL]) in the clinic and home. To make this comparison we will conduct a secondary analysis of data collected in a previous methodological study (AG08947) which examined task performance in five samples of older women including

a subsample of older women with major depression. We chose to analyze the data from the older women subsample because women are often primarily responsible for home managemnt IADL (Rogers et al., 2003) and because they are at a higher risk for major depression than men.

Due to the conflicting evidence in the systematic review described in Chapter 2 of this dissertation, we hypothesize that there will be no difference between the influence of the clinic and the influence of the home on the overall, domain, and task level of performance for our sample.

3.2 METHODS

3.2.1 Participants

Subjects in the original study (AG08947) were recruited from the Benedum Geriatric Center, in Pittsburgh, Pennsylvania. Five different diagnostic groups were recruited. In this data analysis we will be using the data from the *depression group* (n = 59, all females) in which the subjects had a primary diagnosis of major depression, which had remitted (non-psychotic according to the DSM-III R Axis 1) but continued to receive usual treatment only [i.e., medication management in an outpatient clinic]. Also, subjects had to report no other disabling conditions. Additionally, subjects had to: be between 70-84 years of age at time of entry into the study, to report a current limitation in at least one ADL and IADL of concern to the study, to report a prior history of successful task performance in all IADL on the OARS-Multidimensional Functional Assessment Questionnaire [OARS-ADL scale] (Fillenbaum, 1988), reside in the community in a private home or apartment, be able to identify a family member or friend who could serve as a proxy informant, and be medically stable (i.e., had not been hospitalized more than once in the last 3

months for any of the diagnoses and hospitalization was not anticipated within the 3 months following the study.

Subjects were excluded from the study if they had a history of: substance abuse within the past 12 months of the study, central nervous system insults (e.g., stroke or trauma), bipolar affective disorder or psychotic disorder, degenerative disorders of the central nervous system (e.g., Parkinsonism), severe auditory or visual impairment, and/or recently diagnosed cancer.

3.2.2 Instruments

Task performance was assessed with the Performance Assessment of Self-Care Skills [PASS] (Rogers & Holm, 1989). The PASS has two versions: the clinic version (which was used first to assess the subjects' task performance in the clinic) and the home version (which was used 3 days later to assess the subjects' task performance in the home). The PASS is a criterion-referenced assessment tool which contains 26 tasks classified into four functional domains (see Table 3-1): FM (5 items); PC (3 items); CIADL (14 items); and PIADL (4 items). Each task involves a number of subtasks. Performance of each subtask is rated in terms of independence, that is, the assistance needed while performing the task. During task performance, the rater gives assistance (cues) to the person performing the task when needed. The cues range from least assistive (e.g., verbal cues) to most assistive (total assistance) (see Table 3-2). The tasks are rated on the frequency (e.g., occasional or continuous assists) and level (e.g., verbal or physical) of assistance (cues) provided by the examiner. A score from 0-3 is given for each subtask with 0 indicating total assistance in task performance (i.e., person could not perform the task independently) and 3 indicating no assistance needed in task performance (i.e., person performed the task independently) (see Table 3-3). The mean score of all subtasks for an item

yields the task score, the mean score of all tasks in a domain yields the score for that domain, and the mean score of all four domains forms the overall PASS independence score.

The interrater reliability of the PASS was established (Rogers et al., 2003) by administering the PASS to 57 older adults in a clinic (using the clinic version) and home (using the home version) (see Appendix A). Content validity of the PASS was established in reference to other functional assessments; the OARS Multidimensional Functional Assessment questionnaire-Activities of Daily Living (Fillenbaum, 1988; Pfeiffer, 1976), the rating scale for Physical Self-Maintenance and Instrumental Self-Maintenance (Lawton, & Brody, 1969; Lawton, Moss, Fulcomre, & Kleban, 1982), and the Functional Assessment Questionnaire (Pfeiffer, 1987; Pfeiffer, Kurosaki, Harrah, Chance, & Filos, 1982) (see Appendix A).

Table 3-1 PASS domains (4) and tasks (26)

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

Note. From “Disability in Older Adults with Depression” by D. Chisholm, 2005, doctoral dissertation. Copyright 2005 by Denise Chisholm. Adapted with permission.

Table 3-2 PASS prompts hierarchy

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

Note. From "Disability in Older Adults with Depression" by D. Chisholm, 2005, doctoral dissertation. Copyright 2005 by Denise Chisholm. Adapted with permission.

Table 3-3 PASS independence scoring criteria

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

Note. From "Disability in Older Adults with Depression" by D. Chisholm, 2005, doctoral dissertation. Copyright 2005 by Denise Chisholm. Adapted with permission**.**

3.3 DATA ANALYSIS

3.3.1 Participants

Descriptive statistics of the demographic variables were calculated for subjects in 2 datasets: a large depression dataset (n = 444) which was used to anchor the item hierarchy (see item anchoring), and a smaller dataset of clinic-home data of a chosen sample of older women with remitted symptoms of major depression (n = 59). Use of the large dataset of subjects with major depression to anchor the item hierarchy ensured a more valid item difficulty ordering as it contained a wider spread of person abilities. Use of the smaller clinic-home dataset allowed for analyzing differences between clinic and home performance of older women with depression.

SPSS version 15.0 (SPSS Inc, 2002) was used to calculate all descriptive statistics. Participants' descriptive data described the demographic and impairment characteristic.

3.3.2 Data preparation for Rasch analysis

We used Rasch analysis to prepare our data for statistical analysis and examined construct validity and reliability of the PASS using the larger data set. Rasch analysis, developed by George Rasch (1960), is a mathematical method for obtaining objective linear measures of a latent variable (Bond & Fox, 2007). Rasch analysis can also be used to assess the construct validity of measurement scales and assessment tools (Tennant & Conaghan, 2007). It can be conducted using several software packages including WINSTEPS (Linacre, 2006), which is commonly used for moderate- size datasets that contain persons' responses on a set of items. We used WINSTEPS version 3.63.2 (Linacre, 2006) to perform all Rasch analyses for this study. For data preparation, we used WINSTEPS functions of item difficulty estimates, item anchoring, person ability estimates, and item and person hierarchies. For PASS construct validity examination, we used WINSTEPS diagnostic functions of fit statistics, principal component analysis of Rasch residuals, and category functions analysis. For PASS reliability testing, we used WINSTEPS to calculate the item reliability index, person reliability index, person separation index, and internal consistency (Cronbach's alpha). Throughout the METHODS and RESULTS sections we explain in detail these analyses/functions in the following order: data preparation (item difficulty estimate, item anchoring, person ability estimate, and item and person hierarchies); construct validity (fit statistics, principal component analysis and category function analysis); and reliability (item reliability index, person reliability index, person separation index, and internal consistency).

3.3.3 Data preparation: Item difficulty estimates

We first used WINSTEPS to transform the ordinal scores of the PASS into interval scores called logits (log odd units). The logit scale is an interval scale in which the unit interval between the locations on that scale has a consistent value (Bond & Fox, 2007). This transformation process resulted in calculating a logit (i.e., interval score) for each item of the PASS, which determines the relative difficulty of that item compared to the rest of the items of the PASS. The item difficulty is calculated from the total number of persons who passed that item successfully (i.e., scored 3 in the PASS).

3.3.4 Data preparation: Item anchoring

A core requirement of the Rasch model for a valid measurement process is assuring the invariance of item calibrations (i.e., fixing item difficulty estimate for a sample) regardless of the intended purpose for, or the environment of, the measurement (Bond & Fox, 2007). Assuring the invariance of item calibration allows for a valid comparison of items across different samples (males and females) or environments (e.g., clinic and home). To verify item calibration invariance we used WINSTEPS to "anchor" item estimates by using a larger depression database of 444 subjects to create the item hierarchy. Anchored item difficulty logits were then used to establish construct validity and reliability of the PASS.

3.3.5 Data preparation: Person ability estimate

After anchoring the item estimates (i.e., fixing the item hierarchy) we then calculated the person ability estimates for our sample (n = 59) using the anchored item hierarchy. We chose this sample because it was the only subsample of the larger dataset that had clinic and home data for older women with depression. Similar to item difficulty estimation, for each person in the

sample a logit was calculated to indicate the relative ability of each person compared to the rest of the persons in the same sample. The person ability is calculated from the total number of items which the person passed successfully. Person ability logits will be used to compare person overall ability scores between clinic and home, using the paired sample *t*-test.

3.3.6 Data preparation: Item and person hierarchies

Once item difficulty and person ability logits were established, two hierarchies were created: the item difficulty hierarchy (in which items are ranked from the easiest to most difficult) and the person ability hierarchy (in which persons are ranked from the least able/independent to most able/independent). In this study, the midpoint used for item hierarchy was zero. Thus, the greater the logit (more positive) the harder the item is to perform, compared to the rest of the items in the PASS, and the lower the logit (more negative) the easier the item is to perform, compared to the rest of the items in the PASS. In contrast, the greater the logit of a person, the more independent the person is compared to other persons in the sample, and the lower the logit of a person, the less independent the person is compared to other persons in the sample.

3.3.7 PASS construct validity: Fit statistics

Generally, to ensure construct validity, the Rasch model assumes one implicit construct in any measurement process, or unidimensionality. An ideal situation is when the data (or the responses by persons on items targeted to assess their ability) meet the assumption of unidimensionality, or, the real data fit the Rasch model. When the real data fit the Rasch model it indicates that the item and person data relate in a coherent and integrated way that is more likely to represent the construct being assessed (Bond & Fox, 2007). To estimate how well the data meet the unidimensionality assumption, WINSTEPS provides what is called *fit statistics*. Items that do not

fit the model are those that deviate from the expected, linear, unidimensional difficulty pattern. Likewise, persons who do not fit the model are those who deviate from the expected unidimensional ability pattern. As fit statistics are estimations of item/person fit to the Rasch model, they are associated with a degree of error which allow for a more valid interpretation of data misfit. There are two outputs of fit statistics: infit and outfit. Infit statistics are an information-weighted sum (i.e., inlier-sensitive) that detects unexpected responses (misfit/deviated performances) by persons on items targeted at their ability level. In contrast, outfit statistics are outlier-sensitive statistics that detect unexpected responses on items that are very easy or very difficult (Bond & Fox, 2007). Therefore, given that researchers expect people to perform closer to the average ability level, rather than far away from the average ability level, more attention is given to infit statistics than to outfit statistics (Bond & Fox, 2007). Both infit and outfit statistics are the mean of squared residuals (chi square statistics that represent the differences between the model expected values and the real data values) that have an expected value of +1 and a range from 0 to positive infinity. For example, infit mean square values greater than 1, for example, 1.60 (i.e., 1 + 0.60), mean that there is 60% (0.60 *100) more variation in the observed data than the value the Rasch model expected (Bond & Fox, 2007). Outfit mean square values less than 1, for example, 0.80, mean that there is 20% (1 - 0.80*100) less variation in the observed data than the value the Rasch model expected (Bond & Fox, 2007). For clinical functional assessments, such as the PASS, an acceptable mean square error range for infit and outfit statistics is 0.5 to 1.7. Tasks or items with infit or outfit mean square error values greater than the suggested range indicate *noise* (i.e., erratic response), and those less than the suggested range indicate over predictability [i.e., misleading to the notion that we are measuring better than

we really are] (Linacre, 2006). Hence, items with infit and outfit statistics < 0.5 and >1.7 are to be reinvestigated as their misfit may indicate problems in the measurement.

3.3.8 PASS construct validity: Principal component analysis (PCA) of Rasch residuals

Principal component analysis (PCA) of Rasch residuals is a diagnostic analysis that identifies potential deviations from the Rasch core assumption of unidimensionality. It detects "secondary dimensions" in the dataset (different from the main dimension: independence) whose variance was not explained by the Rasch model (Bond & Fox, 2007). The PCA procedure starts with a regular Rasch analysis that constructs a linear (interval) scale of the data, followed by a PCA of the residuals that remained unexplained after the Rasch measure has been extracted. In other words, after performing a Rasch analysis, the remaining unexplained variance in the data is explored (using the PCA) to detect any potential unidentified dimensions in the data that may be affecting the measurement process. The PCA yields empirical and modeled values that represent variances explained and unexplained by the model. The first output of the PCA is an elbow-shaped scree plot which represent the variance log-scale (constructed by Rasch analysis) against the variance components (constructed by the PCA)—the sharper the "elbow" the better the results (less unexplained variance is found by PCA). The second output of the PCA is a group of contrasts (up to five) for the unexplained variances. The largest secondary dimension in PCA is the first contrast, which identifies a common variance of subset items in the residuals. For each contrast, the PCA provides a plot of residual factor loading against the item difficulty logits. The items clustered at the top (positive loading) and at the bottom (negative loading) of the plot represent the most opposing items (i.e., requiring different skills) in that contrast. Usually, those that are clustered at the top have the most substantial loading or effect on the measure. Clustered

items are those to be reconsidered or reinvestigated when assessing the unidimensionality of the instrument (Bond & Fox, 2007; Linacre, 2006).

The PCA also provides another index that is important in investigating the unidimensionality of the measure called the *factor sensitivity ratio* which indicates the substantiality of any additional factor(s) on the stability of the target measure. In other words, it shows how much the factors are responsible for the unexplained variance between the item residuals after the Rasch measure was extracted (Bond &Fox, 2007). The ratio (or the loading) can be calculated by dividing the variance units explained by the first contrast by the variance units explained by the Rasch model. Ratios or loadings of .40 or greater are considered substantial (i.e., substantially affect the stability of the measure).

When interpreting the PCA outputs one should look for the following criteria in the resultant PCA: 1) the variance explained by the Rasch model (should be larger than 60%), 2) the size (eigenvalue) of the unexplained variance explained by the first contrast (should be less than 3.0), 3) the percentage of the unexplained variance explained by the first contrast [should be < 5%], and 4) the factor sensitivity ratio [the smaller the ratio the more compatible the data with the Rasch model] (Linacre, 2006).

It is noteworthy that the PCA of Rasch residuals is different from the common factor analysis. That is, the PCA of Rasch residuals attempts to explain the variance rather than find shared factors (i.e., construct variables). Thus, in the PCA of Rasch residuals "…we want 'not' to find contrasts, and if we do, we want to find the least number of contrasts above the noise level" (Linacre, 2006, p.272) (i.e., the explained variance by the first contrast should be < 5%). In contrast, in factor analysis we search for highly correlated factors that may eventually form a secondary dimension that was not previously identified in the data.

3.3.9 PASS construct validity: Category function analysis

The rating scale of an assessment tool influences the performance of the items in that tool (Bond & Fox, 2007). Thus, the utility of the categories in a rating scale must be empirically tested to ensure that the scale is creating an interpretable outcome (Bond & Fox, 2007). The Rasch model provides a diagnostic application called *category function* in which several characteristics are examined to determine the functionality (or appropriateness) of a rating scale. The category in this context means the responses to an item in an assessment tool (e.g., the PASS has 4 categories for each item; see Table 3-3). The characteristics that are usually examined in a category function analysis are: category frequency, average measures, thresholds, and category fits.

Category frequency is the number of times a category was used/observed across all items for a group of people. The number and distribution shape of category frequency provide a quick basic examination of the rating scale. That is, for a rating scale, the minimum number of responses recommended per a category is 10 and the regular shapes of distribution (e.g., normal or skewed vs. highly skewed distributions) indicate that the categories in that scale are well-utilized. Categories that are infrequently used (usually shaped as highly skewed) are considered either unnecessary or redundant categories (Bond & Fox, 2007).

Average measures are the average of person ability logits of those who chose (scored) a particular category across all items. It is expected that these average measures "monotonically" increase as the scale increases. For example, if the average measure of category 1 in the PASS was 0.8, it would mean that the average person ability for those in the sample who scored 1 is 0.8. Hence, people with higher person abilities would obtain higher category responses and those with lower person abilities would obtain lower category responses.

Thresholds (also known as step calibrations) are "the difficulty estimated for choosing one response category over another" (Bond & Fox, 2007, p. 223-224). Similar to the average measure, thresholds are expected to be monotonic; meaning to increase with an increase in the measured variable (i.e., the harder the category to endorse, the higher the threshold of that category). Thresholds that do not increase monotonically are considered *disordered*. Additionally, the magnitude of the distances between adjoining threshold estimates should not be too close (i.e., categories shadowing or overlapping), nor too far apart (i.e., a large gap between categories), indicating that each threshold represents a distinct position. It is recommended that a threshold should increase by at least 1.4 logits, and not larger than 5 logits, to show distinction between categories. In addition, WINSTEPS provides probability curves (see RESULTS) which aid in inspecting the distance between the thresholds. The curves are graphs that plot the difference between person ability and item difficulty on the x-axis against the probability of choosing a particular response on the y-axis. Each threshold (curve) should have a distinct peak to demonstrate that it is the most probable category for a certain portion of the measured variable. Flat curves represent a redundant or unnecessary category unless the category spans a large portion of the measured variable.

Category fits are infit and outfit mean squares that help in assessing the quality of the rating scale. Outfit values greater than 1.7 indicate that the category is introducing noise into the measurement process and is misinforming rather than informing about the measured variable. Such categories are to be reconsidered. Examining the category function ensures that the ratings reveal responses that are reflective of the construct under investigation (independence).

3.3.10 PASS reliability: Item reliability index

The item reliability index indicates the expected replicability of the item ordering if these items were administered to another sample similar in size and characteristics. Higher item reliabilities indicate greater consistency of item ordering (i.e., item hierarchy). Wider ranges of item difficulty and the larger sample sizes, tend to yield higher item reliability for a sample (Bond & Fox, 2007).

3.3.11 PASS reliability: Person reliability index

The person reliability index indicates the expected stability of person ability hierarchy if these persons were given a similar set of items measuring the same construct (Bond & Fox, 2007). Higher person reliabilities indicate greater consistency of person ordering (person hierarchy). A person reliability index of $\geq$.80 means that the more independent performers can be reliably distinguished from the less independent performers, while a reliability measure of .50 means that the performance differences may be due to random chance. The person reliability index can be affected by a number of factors including: the spread of person ability in the sample, the number of items in the assessment tool, the number of categories per item, and the magnitude of measurement error. Therefore, higher person reliability can be obtained with a wider spread of person ability in a sample, larger numbers of items and categories per item, and lower measurement error.

3.3.12 PASS reliability: Person separation index

The Rasch model provides a person separation index, which helps the researchers examine a tool's ability to consistently differentiate between persons with high abilities and those with low abilities. That is, the person separation index (in logits) indicates the replicability of the

anticipated person ordering if the same set of items (measuring the same construct) is given to another sample. In other words, the higher the index the more reliable and precise is the differentiation between persons of higher ability and those of lower ability, and the more likely this differentiation will be consistent in a similar sample. Values approaching from 2 or higher indicate good person separation index.

3.3.13 PASS reliability: Internal consistency

Cronbach's alpha, indicating the internal consistency of the PASS, was also calculated from raw scores of the sample by WINSTEPS. Values of $\geq$ 0.90 indicate high internal consistency and values between 0.70 - 0.90 indicate moderate reliability (Portney & Watkins, 2000). Both person separation reliability and Cronbach's alpha helped in investigating the reliability of the PASS independence measure.

3.3.14 Performance in the clinic versus home at the PASS overall level

Upon completion of all Rasch analyses we ran paired (or dependent) t-tests to test our hypothesis (the clinic and the home have equal influence on task performance in our sample). We used paired t-tests to compare the sample's overall performance in the clinic and the home. Rasch person ability estimates for each subject were used for this analysis. This helped us understand in which environment the PASS overall level was performed more independently by the persons in our sample.

3.3.15 Performance in the clinic versus home at the PASS domain level

The Rasch model also uses a framework called *differential group functioning* (DGF) which interprets the differences between the person ability of two (or more) groups of persons among a

group of items. To make this differentiation, the DGF uses the person ability logit scores for all subjects in a group, and the item difficulty logits for all items in a group of items. For our study, the person groups will be persons in the clinic and home, and the item groups will be the four PASS domains FM, PC, ADL, and IADL. The DGF will help identify the potential influence of the environment on person ability for the four performance domains. We will conduct four t-tests to examine the difference in person ability logits for FM, PC, CIADL, and PIADL, with Bonferroni corrections. This helped us to understand in which environment the persons in our sample performed tasks more independently, for each domain.

3.3.16 Performance in the clinic versus home at the PASS task level

While some of the domains may not be statistically significantly different between the clinic and home, some of the items may actually show substantially relevant differences between the two environments. Substantial relevance can be described as a substantial difference between two groups in the item difficulty logit that will affect the order of task intervention for a group (Bond & Fox, 2007). These differences are substantially important because regardless of the statistical significance, the differences need to be considered for intervention, as they may provide information about the environment that may not be well highlighted in statistical analyses. Therefore, we explored differences in item difficulty estimates in the clinic and home for all PASS items whether or not the domains showed statistically significant differences between the clinic and home. This helped us to understand which items persons found more difficult in the clinic and in the home. Since the PASS has 26 items we needed to use a powerful method to validly conduct the 26 comparisons. The Rasch model provides a powerful method called Differential Item Functioning (DIF). The Rasch model uses DIF because it requires the relative item difficulty estimate (or item difficulty logit) to be invariant for two similar groups. The DIF

can examine this invariance by determining whether the item difficulty estimate differs significantly between the two groups. First, the item difficulty estimate for all items is calculated and anchored (explained earlier). Then, the two groups are compared against the anchored item hierarchy for all items. For each item, the difference in the DIF measure between two groups is called the *DIF contrast*. The DIF contrast should not be greater than 0.5 logits. A DIF contrast greater than 0.5 logits indicates a potential violation of the invariance of item difficulty estimate between the two groups. It can also indicate items on which groups or classes of persons show substantially relevant differences. WINSTEPS provides statistics to determine the significance of the difference in the DIF contrast. In other words, it determines the probability of observing the DIF contrast when there is no violation of the DIF requirement of invariance. If the probability is significant ($p < .05$), then we can say that the item difficulty estimate is significantly or substantially different between the two groups. If the probability is not significant ($p > .05$), then we can say that the item difficulty estimate is not significantly or substantially different between the two groups.

3.4 RESULTS

3.4.1 Participants

Table 3-4 presents the demographic variables for the large depression dataset (n = 444) which was used to anchor the item hierarchy. Most subjects were females (75.2%) and Caucasian (85.2%). Additionally, most subjects had 12 years of education or higher (81.3%), and only a few were employed when recruited into the study (17.6%). The impairment variables showed that the subjects had mild to moderate depressive symptoms (mean Hamilton Depression Rating scale = 5.83) and low physical impairments (mean Keitel function test = 24.85). Subjects were diagnosed with major depression, which had remitted, and they continued to be treated in an outpatient clinic.

Table 3-4 Demographic and impairment variables for the larger depression dataset

Variables	n = 444
Demographic variables	
Age (years)	74.10
Gender (% females)	75.20
Ethnicity (% Caucasian)	85.20
Marital status (% married)	45.70
Education (% $\geq$ 12 years)	88.60
Occupation (% employed)	17.60
Impairment variables, Mean (SD)*	
HDRS (score range 0 to 14)	5.83 (3.29)
KFT (score range 4 to 100)	26.18 (13.66)

Note. SD = standard deviation; HDRS = Hamilton Depression Rating Scale; KFT = Keitel function test; * = higher scores in HDRS and KFT indicate greater impairment.

Table 3-5 presents the demographic variables for the clinic-home dataset (n = 59) which was used to analyze the differences in the sample's performance between the clinic and the home. All subjects were females with a mean age 75.73 years. The demographic variables showed that the majority of subjects were Caucasian (81.4%), widowed (62.7%), and lived alone (62.7%). Additionally, most subjects had a high school education or higher (81.3%), and had a household income $\geq$ \$10,000 (62.7%). The impairment variables showed that physical impairment was low (KFT), performance speed was moderate (TMT-A, TMT-B), depressive symptoms were minimal to mild (GDS), dementia was not evident ($\geq$78 on 3MS), medical burden (CIRS-G) was low to moderate; and the perceived health performance was moderate. Subjects were diagnosed with major depression, which had remitted, and they continued to be treated in an outpatient clinic.

Table 3-5 Demographics and impairment variables for clinic-home dataset

Variables	*n* = 59
Demographic variables	
Age (years)	75.73
Ethnicity (% Caucasian)	81.40
Marital status (% widowed)	62.70
Living status (% living alone)	62.70
Education (% ≥ high school)	81.30
Household income (% ≥ $10,000)	62.70
Impairment variables, Mean (SD)*	
KFT (score range 4 to 100) (	21.49 (8.51)
TMT-A (scores in seconds)	52.54 (18.17)
TMT-B (scores in seconds)	149.11 (84.71)
GDS (score range 0 to 15)	4.00 (3.75)
3MS (scores range 0 to 100)	93.10 (5.12)
CIRS-G (scores range 0 to 56)	11.57 (3.81)
Perceived health (scores range 0 to 10)	7.39 (1.857)

Note. SD = standard deviation; GDS = Geriatric Depression Scale; KFT = Keitel Functional Test; TMT-A = Trail Making Test-A; TMT-B = Trail Making Test-B; 3MS = Modified Mini-Mental State Examination; CIRS-G = Cumulative Illness Rating Scale for Geriatrics. * = higher scores in KFT, TMT-A, TMT-B, GDS, and CIRS-G indicate greater impairment, and lower scores in 3MS and perceived health status indicate less impairment.

3.4.2 Data preparation: Item difficulty estimates

All PASS clinic and home raw scores were transformed into logits which indicated the item relative difficulty. Item difficulty logits were then used to anchor the item difficulty hierarchy (see Table 3-6).

3.4.3 Data preparation: Item anchoring

The item hierarchy was anchored using the larger depression dataset (n = 444). Trimming toenails was the most difficult among the 26 items, and indoor walking was the easiest (See Table 3-6).

3.4.4 Data preparation: Person ability estimate

Based on the anchored item hierarchy, person ability was calculated for each person in both the clinic and home. Person ability estimates were then used to run paired t-tests to compare the overall ability of each person in the clinic and the home. Figure 3-1 illustrates the distribution of person ability in the clinic and home. Figure 3-2 presents the item-person map which illustrates the person ability of each person in the sample in relation to the anchored item difficulty.

3.4.5 Data preparation: Item and person hierarchies

The item hierarchy was first created and anchored (See Table 3-6). Then, person ability hierarchy was created which allowed for comparing the performance of all 59 subjects in the clinic and home at the overall level (Figure 3-1 and Figure 3-2).

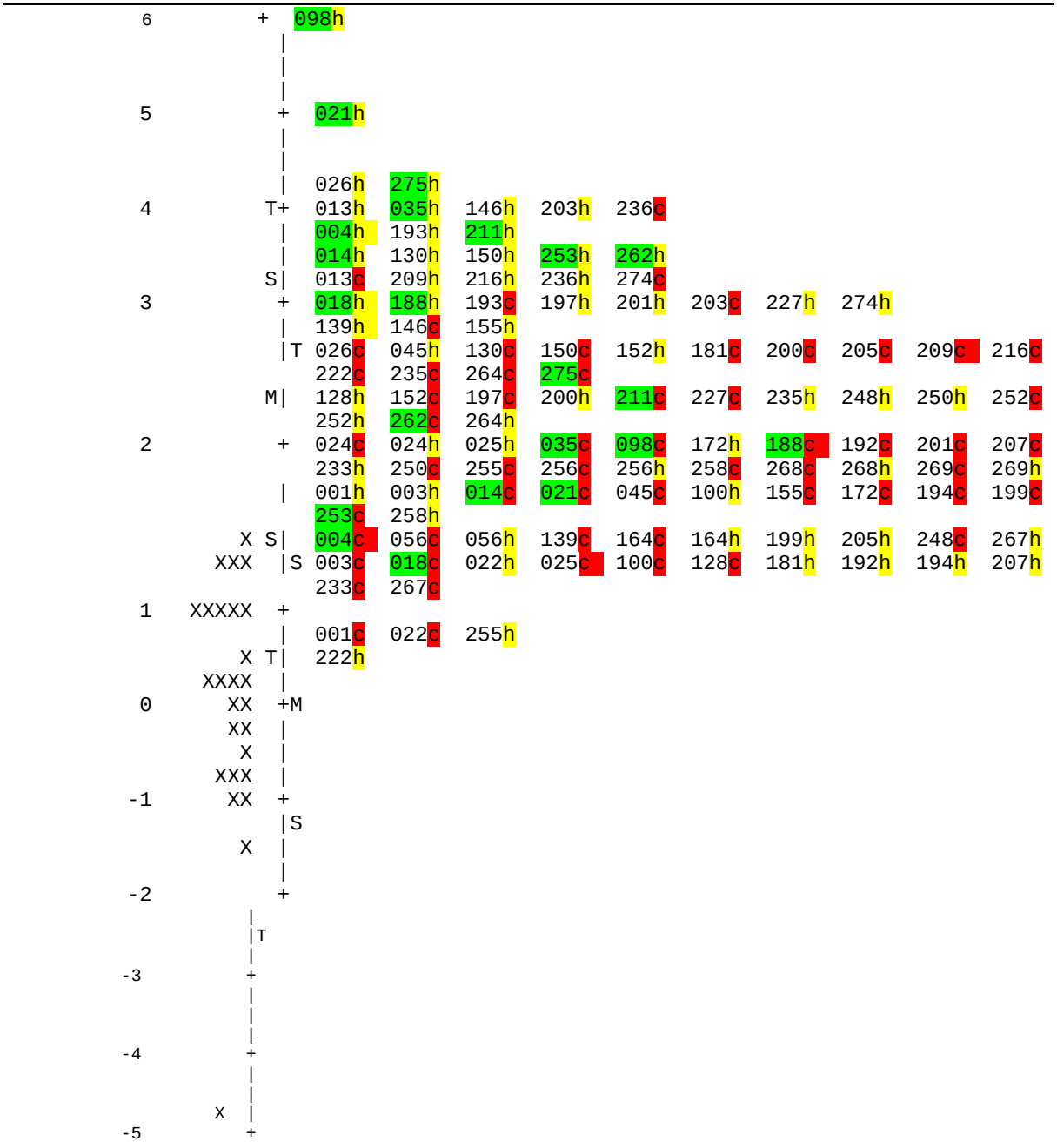
Table 3-6 Rasch item difficulty logits and fit statistics of the PASS

Pass Tasks	Logits	INFIT MNSQ	OUTFIT MNSQ	Raw* scores
Most difficult				
Trim toenails	1.39	1.76	1.88	512
Shop	1.21	0.46	0.74	915
Change bed linens	1.15	1.03	0.91	549
Balance a checkbook	1.13	0.57	0.66	898
Mail bills and checks	1.10	0.51	0.61	926
Stovetop use	1.01	0.97	0.96	932
Oven use	0.99	0.77	0.69	543
Pay bills by check	0.94	0.57	0.65	970
Medication management	0.91	0.56	0.77	997
Bathtub/shower transfer	0.43	1.56	1.44	1091
Use sharp utensils	0.37	1.14	0.90	1072
Small repairs	0.36	0.84	0.82	1103
Home safety	0.31	0.75	0.91	1111
Clean up after meal preparation	0.25	1.65	1.20	1082
Telephone use	0.05	1.00	1.03	1156
Stair use	0.04	2.59	1.98	1099
Obtain information: visual	-0.14	1.30	1.17	1176
Bending, lifting, carrying	-0.33	1.74	1.42	1196
Obtain information: auditory	-0.54	1.17	1.02	1215
Dress	-0.79	1.15	0.91	1237
Sweep	-0.85	2.33	1.38	1241
Oral hygiene	-0.86	2.06	1.58	1236
Play bingo	-0.88	1.77	0.94	1201
Bed transfer	-0.97	1.34	0.97	1243
Toilet transfer	-1.50	1.66	0.99	1275
Indoor walking	-4.79	1.03	0.74	1318
Least difficult				
Mean (SD)	0.00 (1.25)			

Note. Measure = item difficulty logits; MNSQ = mean square; * = Raw scores calculated by the Rasch model as the sum of the scored responses to an item.

Best Performance (Most Independent)		
Logit values for clinic (n =59)		Logit values for home (n = 59)
	6.00	6.35
		5.11
	5.00	
	4.00	4.36; 4.36
3.89		3.81; 3.81; 3.86; 3.89; 3.89; 3.89; 3.89
	3.50	3.54; 3.54; 3.54; 3.54; 3.54
		3.25; 3.25
3.25; 3.25; 3.01; 3.01	3.00	3.01; 3.01; 3.01; 3.01; 3.17
2.79		2.79; 2.79; 2.92; 2.97
2.59; 2.59; 2.59; 2.59	2.50	2.59
2.41; 2.41; 2.41; 2.41; 2.41; 2.41; 2.41; 2.41		2.36; 2.41
2.24; 2.24; 2.24; 2.24; 2.24; 2.24		2.24; 2.24; 2.24; 2.24; 2.24; 2.24
2.09; 2.09; 2.09; 2.09; 2.09	2.00	2.04; 2.09; 2.09; 2.09
1.94; 1.94; 1.94; 1.94; 1.94; 1.94; 1.94; 1.94; 1.81; 1.81; 1.81; 1.81; 1.81		1.81; 1.89; 1.94; 1.94
1.68; 1.68; 1.68; 1.57; 1.57; 1.57; 1.57	1.50	1.51; 1.57; 1.57; 1.67; 1.68; 1.68
1.46; 1.36; 1.36; 1.36		1.36; 1.46; 1.46
1. 1.17; 1.17; 1.17	1.00	1.17; 1.17; 1.26; 1.26
0.79; 0.79		0.86
	0.00	0.53

Figure 3-1 Person ability estimates (in logits) for persons in the clinic and home



Note. Item-person map illustrates the person ability (right side) for persons in the clinic (c) and home (h) in relation to the anchored item difficulty hierarchy (left side). Refer to Table 3-6 for item difficulty hierarchy;* Yellow = persons in the home; Red = persons in the home; Green = persons with large difference in their person ability logits between the clinic and home (greater ability at home).

Figure 3-2 Item-person map illustrates the person ability

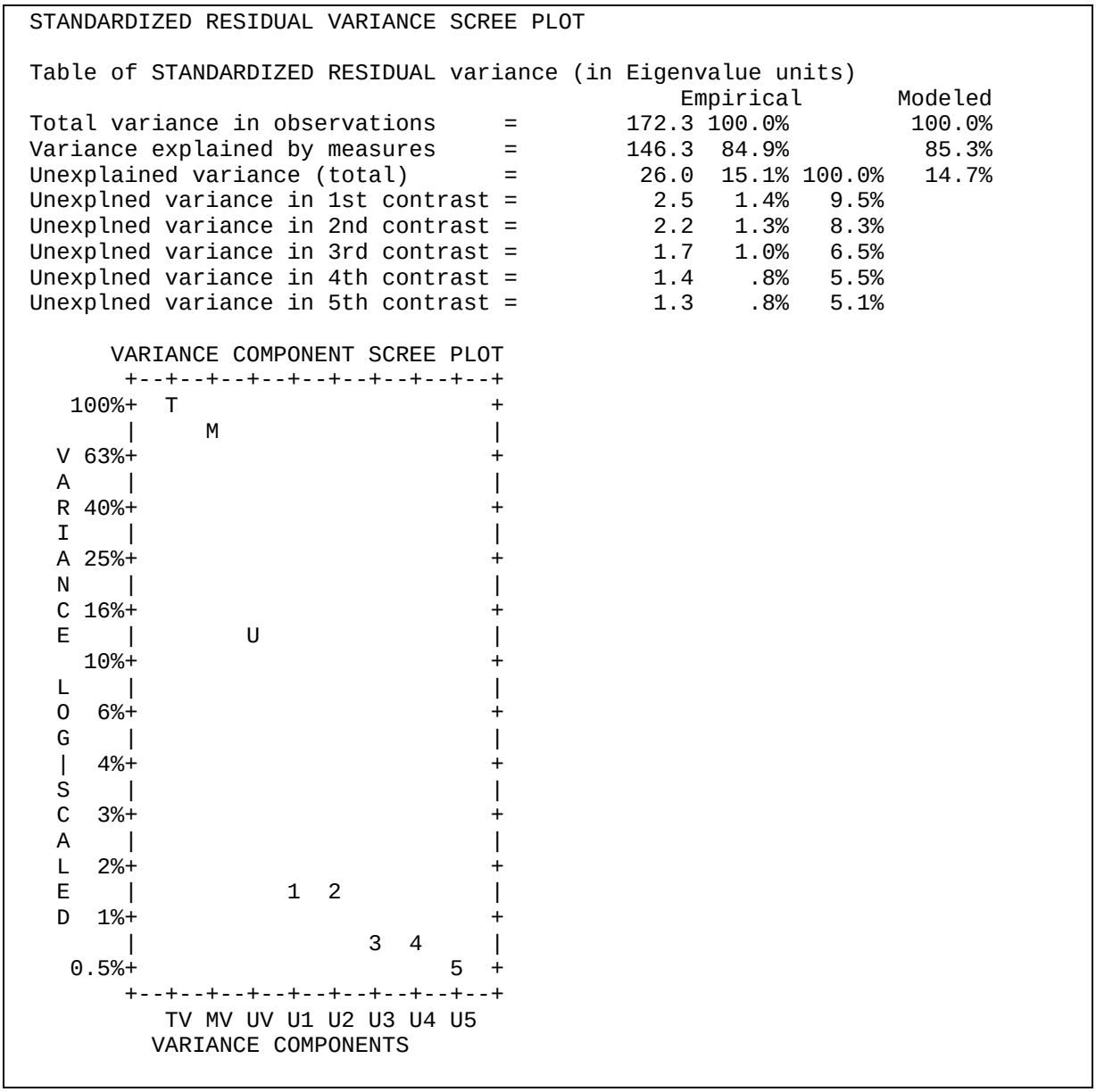
3.4.6 PASS construct validity: Fit statistics

Table 3-6 summarizes the anchored item difficulty logits and fit statistics of PASS items for our sample. The table shows that except for the tasks trim toenails and stair use, PASS tasks were compatible with the Rasch model as the majority of the items had either one or both fit statistics within the acceptable range ($\leq 0.5 \geq 1.7$).

3.4.7 PASS construct validity: Principal component analysis (PCA) of Rasch residuals

The PCA for the PASS is presented in Figure 3-3. The Rasch model explained 84.9 % of the total variance (100%). If the data fit the Rasch model perfectly, the measurement dimension would explain 85.3% of the variance. The unexplained variance for PASS was 15.1%, and the first contrast in the residuals explained 1.4% of the variance. The eigenvalue of the first contrast was 2.5, indicating that it had the strength of about 2 or 3 items out of 26 items in this analysis. Hence, based on the previously explained criteria of PCA, the variance explained by the first contrast is less than 5% (i.e., noise) which indicates that "independence" in the PASS (or the underling construct of amount of assistance provided), for our sample, was unidimensional. PCA item loadings are presented in Table 3-7. Figure 3-4 visually illustrates the factor loading against the item difficulty logits for each task in the PASS. The residuals appear to have a random spread, although the opposed poles (i.e., the items at the top of the plot and those at the bottom) may represent common items (i.e., potential secondary dimensions). For example, the "A" (mailing bills and checks) and "a" (clean up after meal preparation) in the plot represent the items in the PASS that had the largest positive (A) and negative (a) loadings in the first contrast of the residuals (i.e., most opposed items). As the PCA is "indicative" rather than "definitive" about secondary dimensions, "A" and "a" may require different abilities rather than opposed dimensions (e.g., cognitive vs. physical abilities). Finally, the results showed that the factor

sensitivity ratio was 1.7% (2.5/146.3 =.017), considerably less then .40, indicating that the variance of items explained in the first contrast do not substantially affect the stability of the measure. All in all, the PCA showed that the PASS construct of independence is unidimensional.



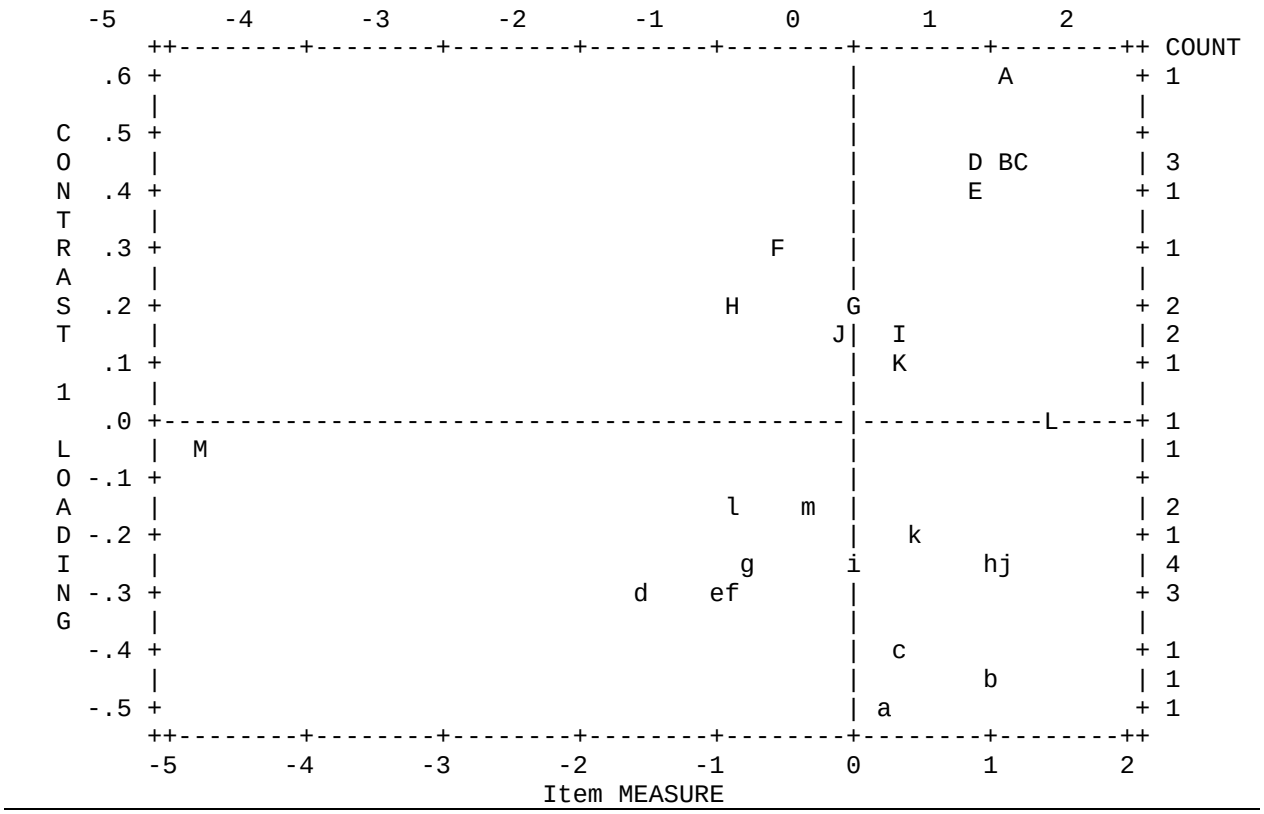
Note. T, TV represents the total variance in the observations. M, MV represents variance in the observations explained by the Rasch measures. U, UV represents unexplained variance. 1 (U1) represent first contrast (component) in the residuals. 2 (U2), 3 (U3), 4 (U4) and 5 (U5) represent second, third,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

Figure 3-3 Standard residual variance scree plot of principal components analysis (PCA) for Performance Assessment of Self-care Skills (PASS)

Table 3-7 Matrix of standardized residual contrast 1 of principal components analysis (PCA) for the PASS

	PASS tasks	Loading	Measure
A	Mailing bills and checks	0.58	1.10
B	Balancing a checkbook	0.47	1.13
C	Shopping	0.44	1.21
D	Medication management	0.43	0.91
E	Paying bills by check	0.40	0.94
F	Obtaining information: auditory	0.30	-0.54
G	Telephone use	0.22	0.05
H	Playing bingo	0.18	-0.88
I	Home safety	0.17	0.31
J	Obtaining information: visual	0.17	-0.14
K	Small repairs	0.12	0.36
L	Trimming toenails	0.00	1.39
a	Cleanup after meal preparation	-0.48	0.25
b	Stovetop use	-0.43	1.01
c	Use sharp utensils	-0.39	0.37
d	Toilet transfer	-0.31	-1.50
e	Bed transfer	-0.29	0.97
f	Sweeping	-0.28	0.85
g	Dressing	-0.26	0.79
h	Oven Use	-0.24	0.99
i	Stair use	-0.24	0.04
j	Changing bed linens	-0.23	1.15
k	Bathtub/shower transfer	-0.20	0.43
l	Oral hygiene	-0.15	-0.86
m	Bending, lifting, carrying	-0.13	-0.33
M	Indoor walking	-0.06	-4.79

Note. Measure = item difficulty logits.



Note. The plot shows the first contrast's loadings and item measures (item difficulty logits) in the matrix of standardized residual contrast 1. This plot presents the clustered residuals visually.

Figure 3-4 Standardized residual contrast 1 plot of principal components analysis (PCA) for Performance Assessment of Self-care Skills (PASS) PASS construct validity

Table 3-8 represents the category function analysis of the PASS. The *observed percentage* column represents the frequency for each PASS scoring category, and it was least observed for category "1" (occasional physical assists to continuous verbal assists). The *observed average* column represents the average measure for the categories which increased in size as the variable increased. This indicated that, on average, the persons with higher abilities (more independent) favored the higher category (scored "3" in the PASS). The *structure calibrations* column represents the thresholds of categories. Category 1 was the most difficult to be observed. The distance between the thresholds of the adjacent categories was almost within

the recommended range of 1.4 to 5 logits. The probability curves for the categories thresholds are illustrated in Figure 3-5. The curves show that each category (except 1) had a distinct peak which indicates that each category represents (measure) a distinct level of ability. Although category 1 appeared to be flat (no peak), it is still a useful category as it spanned a large portion of the construct. The column of *INFIT MNSQ* and *OUTFIT MNSQ* show fit statistics for each category. All fit statistics were within the acceptable range, less than 2, indicating that there was no noise introduced into the measurement process. In summary, category function analysis showed that the categories of the PASS were functioning appropriately.

Table 3-8 Category function of the PASS

Category	Observed %	Observed average	Structure calibration	INFIT MNSQ	OUTFIT MNSQ
0	5	-0.01	NONE	1.21	1.77
1	1	0.68	1.36	1.28	1.62
2	25	0.99	-2.12	0.88	0.73
3	62	2.47	0.76	0.98	0.98

Note. MNSQ = mean square.

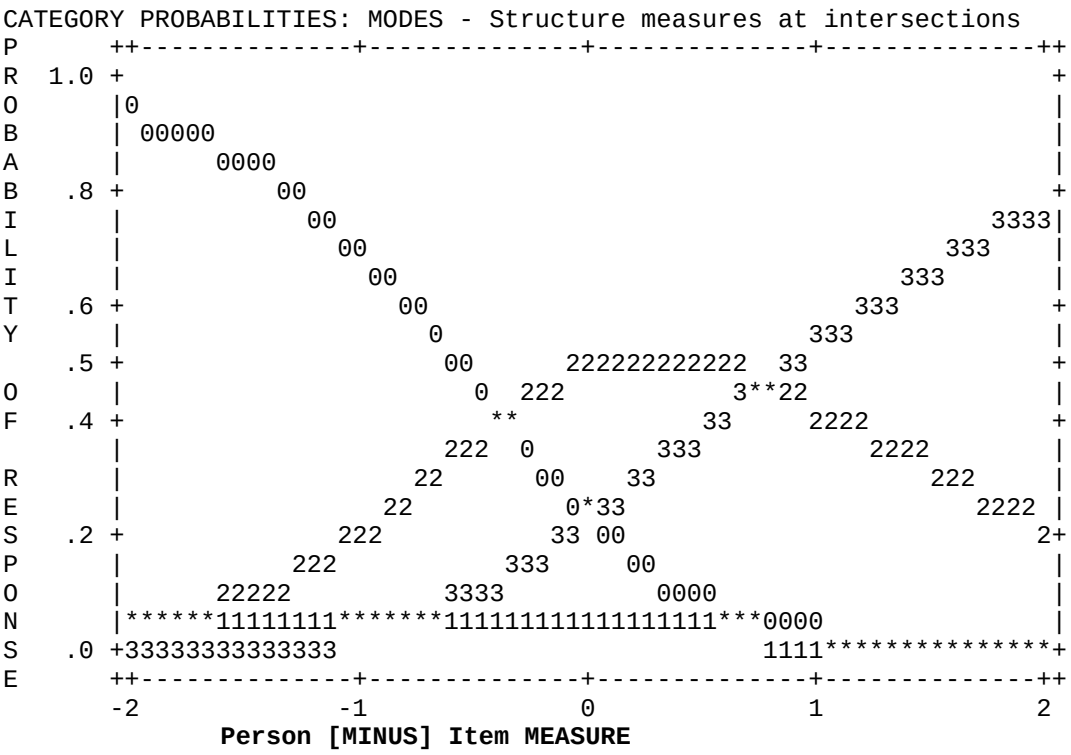


Figure 3-5 Probability curves for the PASS categories

3.4.8 PASS reliability: Item reliability index

The item reliability index was 0.98 indicating excellent replicability of the item hierarchy, or that item location on the difficulty log scale would be consistent when the same set of items are used with other samples of similar traits and size.

3.4.9 PASS reliability: Person reliability index

The person reliability index was 0.76, indicating good replicability of the person ordering if a similar set of items, measuring the same construct, was used with this sample.

3.4.10 PASS reliability: Person separation index

The person separation index was 1.76 which indicates that the PASS adequately separates persons with higher abilities from those with lower abilities.

3.4.11 PASS reliability: Internal consistency

The internal consistency of the PASS was also excellent, with a Cronbach's alpha of 0.90.

3.4.12 Performance in the clinic versus home at the PASS overall level

Table 3-9 shows the *t*-test results. The person ability logits indicated that persons were significantly more independent in the home than in the clinic for overall PASS performance ($p <$ .001).

Table 3-9 Comparison between clinic and home overall PASS person ability scores

Environment	Mean*	*t*	*df*	*SE*	Significance
Clinic	2.04				
		-4.24	58	0.137	p < .001
Home	2.63				

Note. *df* = degrees of freedom; *SE* = standard error; * = mean person ability logit scores for each group.

3.4.13 Performance in the clinic versus home at the PASS domain level

Table 3-10 summarizes the differences in person ability across the four domains in the clinic and home. After applying Bonferroni corrections, the person ability was not significantly different between clinic and home for FM and PC (where home was better for both domains). However, person ability was significantly different in the clinic and home for CIADL and PIADL. The home was significantly more facilitative than the clinic for both domains ($p < 0.001$). Figure 3-2

showed that when tasks were performed at home more subjects showed higher ability (more independence) than when tasks were performed in the clinic.

Table 3-10 Person ability differences between the clinic and home for the PASS domains.

PASS domains	Environment	Mean*	*t*	*df*	Significance **
FM	Clinic Home	2.06 2.07	-0.05	58	p <0.96
PC	Clinic Home	2.15 2.21	-0.33	58	p <0.75
CIADL	Clinic Home	2.68 3.49	-5.26	58	p <0.001
PIADL	Clinic Home	2.60 3.32	-3.45	58	p <0.001

Note. FM = functional mobility; PC = personal care activities of daily living; CIADL = cognitive-oriented instrumental activities of daily living; PIADL = physical -oriented instrumental activities of daily living; *df* = degrees of freedom; * = person ability logits; ** significant at $p < 0.013$ with Bonferroni corrections.

3.4.14 Performance in the clinic versus the home at the PASS task level

Table 3-11 summarizes some of the DIF results. Performance was statistically significantly more independent (less difficult) in the home for dressing in the PC domain, for mail bills and checks, telephone use, and obtaining information: auditory in the CIADL domain and for clean up after meal preparation in the PIADL domain. Performance was (substantially relevant [DIF contrast greater than 0.5] but not statistically significant [p > .05]) more independent in the home for toilet transfer, bending, lifting, carrying, and playing bingo. Performance was significantly more independent in the clinic for stair use, and bathtub/shower transfer in the FM domain, for trimming toenails in the PC domain, and for oven use in the CIADL domain. Finally,

performance was (substantially relevant but not statistically significant) more independent in the clinic for shopping. Figure 3-6 visually illustrates the DIF measures for the clinic and home. Although the figure shows differences in item difficulty estimates for several items, the significance of these differences can only be determined from the Table 3-11. Figure 3-7 summarizes the results on the differences between the clinic and home at the overall, domain, and task levels.

Table 3-11 Differences in person ability in the clinic and home at the task level

PASS tasks	DIF contrast	Probability	Less difficult in
Bed transfer	0.30	p <0.51	Clinic
Stair use	2.39	p <0.00*	Clinic
Toilet transfer	0.65	p <0.19	Home
Oral hygiene	0.19	p <0.69	Home
Bathtub/shower transfer	0.59	p <0.04*	Clinic
Trim toenails	1.43	p <0.00*	Clinic
Dress	1.20	p <0.00*	Home
Shop	0.60	p <0.81	Clinic
Pay bills by checks	0.44	p <0.15	Clinic
Balance a checkbook	0.04	p <0.88	Home
Mail bills and checks	0.62	p <0.05*	Home
Bending, lifting carrying	0.81	p <0.08	Home
Telephone use	2.03	p <0.00*	Home
Medication management	0.04	p <0.90	Home
Change bed linens	0.06	p <0.85	Clinic
Obtaining information: auditory	1.29	p <0.04*	Home
Obtaining information: visual	0.02	p <0.98	Clinic
Small repairs	0.21	p <0.52	Home
Sweep	0.16	p <0.81	Home
Indoor walking	1.15	p <0.61	Home
Home safety	0.13	p <0.69	Home
Play bingo	0.95	p <0.18	Home
Oven use	0.66	p <0.05*	Clinic
Stovetop use	0.12	p <0.78	Clinic
Use sharp utensils	0.44	p <0.23	Home
Clean up after meal preparation	1.05	p <0.04*	Home

Note. DIF = differential item functioning; * = item difficulty estimate is significantly different in the clinic and home for that item.

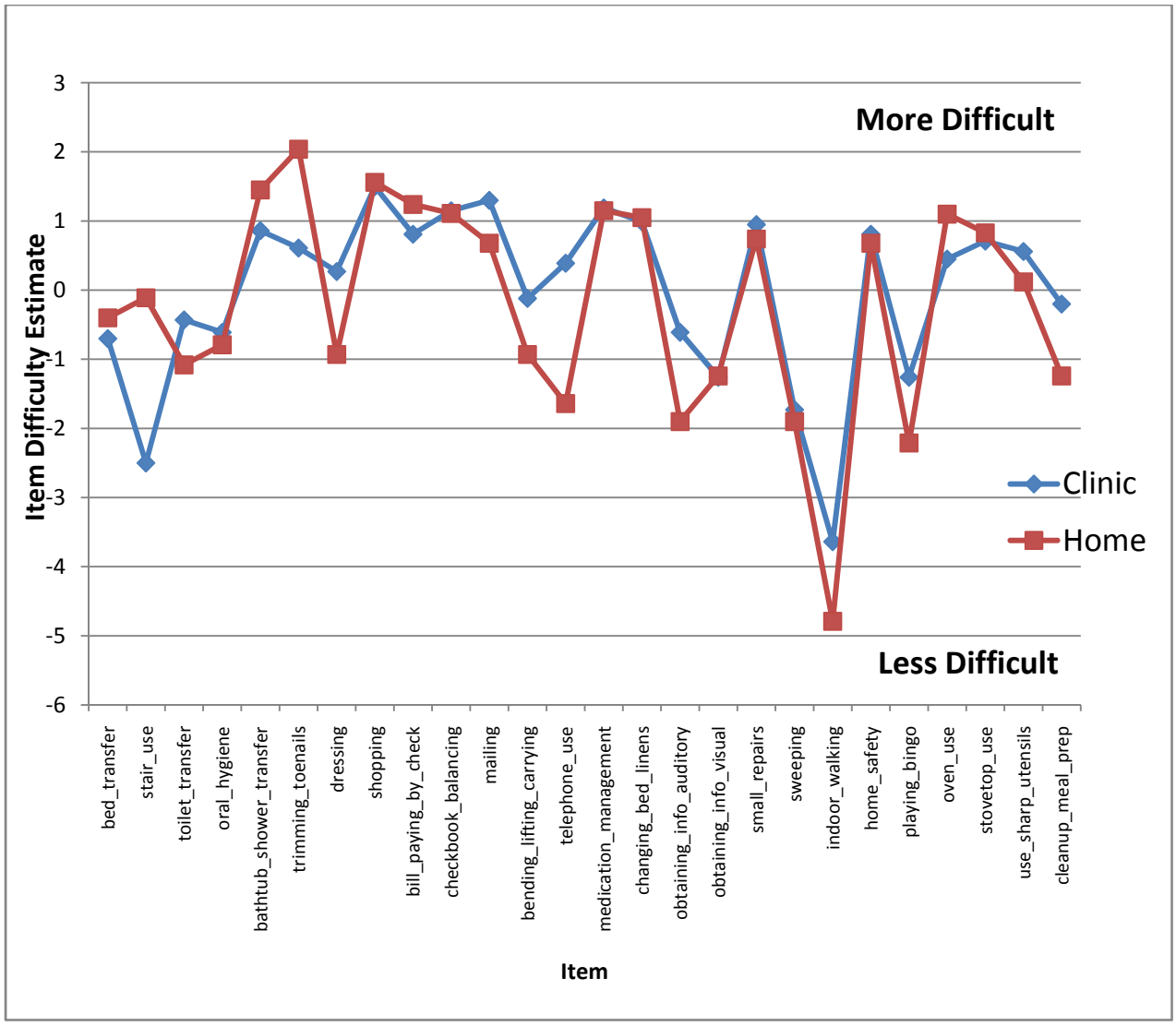


Figure 3-6 Comparison between the item difficulty estimates in the clinic and home

OVERALL LEVEL

Home > Clinic

DOMAIN LEVEL

Home> Clinic

FM PC CIADL* PIADL*

TASK LEVEL

Home > Clinic

FM	*PC*	*CIADL*	*PIADL*
Toilet transfer*	Oral hygiene	Balance a checkbook	Bending, lifting, carrying*
Indoor walking	Dressing **	Mail bills**	Sweeping
		Telephone use **	Clean up after meal preparation **
		Medication management	
		Obtain info: auditory **	
		Small repairs	
		Home safety	
		Play bingo*	
		Use sharp utensils	

Clinic > Home

FM	*PC*	*CIADL*	*PIADL*
Bed transfer	Trimming toenails	Obtaining info: visual	Change bed linens
Stair use **		Oven use**	
		Shopping*	
Bathtub/ shower transfer **		Pay bills by checks	
		Stovetop use	

Note. FM = functional mobility; PC = personal care activities of daily living; CIADL = cognitive-oriented instrumental activities of daily living; PIADL = physical-oriented instrumental activities of daily living. > = performance more independent in; ** = statistically significant and substantially relevant at $p < .013$ with Bonferroni corrections; * = substantially relevant but not statistically significant ($p > .05$).

Figure 3-7 Summary of all statistical (**) and substantial (*) differences between the clinic and home at the overall, domain and task levels of the PASS.

3.5 DISCUSSION

Every year a large number of older adults, mostly older women, will be diagnosed with major depression. Many of those older adults will encounter difficulties in ADL and IADL (National Institute of Mental Health, 2007). Environment is one of several factors that may either significantly increase or decrease these difficulties, and different environments influence these difficulties differently. Despite the complicated nature of major depression in late life and the critical role that may be played by the environment on task performance of ADL and IADL, the influence of the environment on ADL and IADL for this population has been ignored in past research. Our findings indicated that the environment does affect the performance of ADL and IADL at several levels for older women with major depression. We rejected the hypothesis for overall performance because performance in the home was significantly better than performance in the clinic, as it was for the CIADL and PIADL domains, and for several tasks. Clinically, understanding the influence of environment on the performance of everyday tasks potentiates the ability of practitioners to minimize environmental challenges and facilitate independence in ADL and IADL necessary for community living.

The main findings of the study indicate that environment significantly affects overall performance of ADL and IADL, with performance being more independent in the home. Bearing in mind that the samples in previous studies were different, this finding was in contrast to other studies that compared differences in overall performance between the clinic and home for older women with heart failure (Raina, Rogers, & Holm, 2007), older women with osteoarthritis of the knee (Rogers et al., 2003), and older adults with stroke (Sheikh et al., 1979) (see Table 3-12 and Table 2-1 in Chapter 2). The item-person map helps to

visually illustrate this finding by showing that person ability at home was generally greater than in the clinic for most of the subjects. In other words, more persons had the ability to perform more difficult items in the home than they did in the clinic. One possible explanation for these results is that the familiarity of the home can facilitate or enhance a more independent overall performance of daily activities, especially for persons who lack the motivation to explore new environments or even go out of the home. This familiarity is inherently absent in clinic facilities where performance assessment is usually conducted instead of, or prior to, a performance assessment at home. The important implication of this significant difference is to question the trend adopted by healthcare professionals during discharge planning of predicting functional task performance at home based on the outcomes of a functional assessment in the clinic. Such predictions may result in important decisions that directly or indirectly affect the quality of life of the patient being assessed. Our findings indicate that overall, independence in ADL and IADL was negatively impacted in the clinic environment. Moreover, this is "usual care." With major depression, if patients are even evaluated for ADL and IADL the assessment is usually conducted in a clinic setting, leading to an underestimation of performance compared to the "lived in" environment. This, in turn, can lead to inappropriate recommendations being made, which affect the quality of life of that patient on the functional, familial, financial, and social levels. More importantly, for people with major depression, who have affective symptoms such as the loss of motivation, apathy, and depressed mood; somatic symptoms such as fatigue, psychomotor retardation, loss of energy, and restlessness; and cognitive symptoms such as difficulties in decision making, thinking, and concentration (American Psychiatric Association, 2000) the challenges presented by an unfamiliar clinic environment may

present an assessment burden that is beyond the coping abilities of the patient. Our findings suggest that it is more appropriate to assess functional performance in the home -- it is familiar, and is ultimately the "lived in" environment. Thus, there is no need to infer independence in performance from one environment to another, namely from the clinic to the home. According to our results, assessment in the clinic of the overall performance of women with late life depression should be viewed with caution, because it may underestimate performance.

Consistent with the facilitative nature of the home on overall ADL independence, our sample performed CIADL and PIADL domain items significantly more independently in the home than in the clinic. In comparison to the findings of the previous studies that assessed performance at the domain level, our finding about the CIADL was consistent with Raina, Rogers, and Holm (2007), but not with Rogers et al. (2003). However, our finding about the PIADL was not consistent with either of the two studies (see Table 3-12 and Table 2-1 in Chapter 2). A possible explanation of these findings is that the items in these two domains have a cognitive (attention, problem-solving, processing speed) or physical (lifting and transferring objects, bending, carrying and moving oneself across rooms) emphasis which may be more enhanced by a familiar environment such as the home. For example, the task clean up after meal preparation (PIADL) may be easier to perform, and take much less effort in a familiar environment, because the subject is familiar with the physical space and the location of the cleaning tools. Further, although the differences between the performance in the clinic and home for the FM and PC domains were not statistically significant, the fact that the persons in the sample performed the items of these two domains more

independently in the home than in the clinic, further supports the notion of the facilitative influence of a familiar, "lived in" environment for older women with late life depression.

Finally, at the item level, Rasch analysis offered further support that performance was more independent in the home for many tasks. That is, the home was statistically more facilitating than the clinic for dressing, mail bills and checks, telephone use, obtaining information: auditory, and clean up after meal preparation. Moreover, home was more facilitating (but not statistically significant) than the clinic for toilet transfer, bending, lifting, carrying, and playing bingo. Our results at the task level were consistent with Raina, Rogers, and Holm (2007) (which is the only study that used the PASS and assessed performance at the task level) on some of the items in FM (stair use), PC (trimming toenails), CIADL (telephone use, home safety, and small repairs), and PIADL (clean up after meal preparation). However, our results were not consistent with Raina et al. regarding the remaining PASS tasks (see Table 3-12 and Table 2-1 in Chapter 2). The performance of most, if not all, of these items can be enhanced by a familiar environment that reduces extra environmental distracters (novelty of the space) and directs the attention toward the task being performed. When all attention is directed toward the task, fewer errors would likely be observed, and less help would be needed in performing the task. In additions, the clinic was more facilitative than the home for stair use, bathtub/shower transfer, trimming toenails, and oven use. All these items require tangible, facilitating, standardized features in the performance environment such as clutter-free spaces and good lighting rather than familiarity to successfully perform the items. Although some of the findings did not reach statistical significance, the substantial relevance of these differences calls for extra attention when assessing older women with late life depression.

Our study was the first to use a sample of older women diagnosed with major depression to examine the influence of the environment on task performance. All previous studies (see Chapter 2) used mixed samples of older adults (males and females) and focused on physical impairments. Our aim was to identify the influence of the clinic and home at the overall, domain, and task levels of performance. Since the literature did not provide evidence in this regard (see Chapter 2) we hypothesized that there would be no difference between the influence of the clinic and home at the overall, domain, or task levels of performance. To achieve our aim, we compared the performance of 59 older women with major depression in the clinic and home using the PASS, which is a performance-based, criterion-referenced assessment tool. Based on traditional psychometric properties, the reliability and validity of the PASS have been established in several previous studies (see Appendix A). Nonetheless, to be more confident of the appropriateness of the Rasch model with a psychosocial diagnostic sample, namely major depression, we further investigated the validity and reliability of the PASS. For that purpose, we used several diagnostic functions in the Rasch model (explained earlier). The results of these diagnostic functions ensured the validity and reliability of the PASS. For example, the excellent internal consistency measure, the compatible fit statistics, and the PCA, all confirmed the unidimensionality of the PASS construct of independence. In addition, the category function analysis confirmed the construct validity of the PASS and established that the rating categories of the PASS were highly appropriate and reflective of different performance levels. Finally, the tests of item reliability, person reliability, and the person separation index confirmed the stability of the PASS for consistently differentiating between difficult and easy items, and between more independent and less independent persons. All in all, the above mentioned diagnostic

functions confirmed that the PASS is a highly valid and reliable assessment tool to evaluate ADL and IADL performance in older women with major depression.

After ensuring the validity and reliability of the PASS we used the Rasch model to compare ADL and IADL performance in the clinic and home, by converting the PASS ordinal scores into interval scores. The use of the anchored item hierarchy against which the persons' performances in the clinic and home were compared, made the comparison even more valid. Additionally, the use of objective item difficulty and person ability estimates calculated with the Rasch model allowed us to use these estimates rather than the raw scores in other statistical tests such as t-tests and DIF. Hence, the excellent validity and reliability of the PASS and the powerful objectivity of the Rasch analysis allowed for a more valid comparison of the sample's performance in the clinic and the home, thereby increasing the likelihood of generalizing our results to other similar populations.

The use of a sample of older women limited the generalization of our results to other samples of men and women, or of men only, with major depression. Hence, future studies that use samples of men and women, or samples of men only, are needed to obtain a clearer picture of the influence of the environment on task performance in older adults with major depression. Moreover, our study examined the influence of two environments only, the clinic and home, on task performance. Thus, future studies that examine the influence of other environments that impose additional demands on task performance are needed for better understanding of this influence. The additional demands may include social interaction which can be challenging for older adults with major depression (e.g., shopping tasks in the community instead of simulated shopping tasks) or using public facilities that may not be as comfortable and private as those at home (e.g., toilet use in public restrooms).

In summary, the results of our study showed that the environment may actually affect task performance on several levels. This calls for the attention of functional performance assessors in clinical practice. The familiarity of the home seemed to facilitate the overall functional performance, and the performance of the IADL functional domains, for older women with major depression. In contrast, the "non standardization" of the home and the novelty of the clinic seemed to hinder some of the tasks. The implications of our results may help practitioners to direct their functional assessment focus to those domains and items which are known to be more difficult, based on the assessment setting. Hopefully our findings will also be used to support the need for functional assessment in the home for older women with major depression. Because assessment data are used to guide intervention, assessment in the "lived in" environment is more efficient and accurate for addressing ADL disability interventions for older women with late life depression who wish to remain in the community.

Table 3-12 Consistency of our results with the literature

Performance level	Our results (performance more independent in)	Consistent with	Not consistent with
Overall	Home		Raina, Rogers, & Holm (2007) Rogers et al., (2003) Sheikh et al. (1979)
Domain			
FM	Home		Raina, Rogers, & Holm (2007) Rogers et al., (2003)
PC	Home		Raina, Rogers, & Holm (2007) Rogers et al., (2003)
CIADL	Home*	Raina, Rogers, & Holm (2007)	Rogers et al., (2003)
PIADL	Home*	Raina, Rogers, & Holm (2007)	Rogers et al., (2003)
Task *			
Bed transfer	Clinic		Raina, Rogers, & Holm (2007)
Stair use	Clinic ¥	Raina, Rogers, & Holm (2007)	
Toilet transfer	Home $		Raina, Rogers, & Holm (2007)
Oral hygiene	Home		Raina, Rogers, & Holm (2007)
Bathtub/shower transfer	Clinic ¥		Raina, Rogers, & Holm (2007)
Trim toenails	Clinic ¥	Raina, Rogers, & Holm (2007)	
Dress	Home ¥		Raina, Rogers, & Holm (2007)
Shop	Clinic $		Raina, Rogers, & Holm (2007)
Pay bills by checks	Clinic		Raina, Rogers, & Holm (2007)
Balance a checkbook	Home		Raina, Rogers, & Holm (2007)
Mail bills and checks	Home ¥		Raina, Rogers, & Holm (2007)
Bending, lifting carrying	Home $		Raina, Rogers, & Holm (2007)
Telephone use	Home ¥	Raina, Rogers, & Holm (2007)	
Medication management	Home		Raina, Rogers, & Holm (2007)
Change bed linens	Clinic		Raina, Rogers, & Holm (2007)
Obtaining information: auditory	Home ¥		Raina, Rogers, & Holm (2007)
Obtaining information: visual	Clinic		Raina, Rogers, & Holm (2007)
Small repairs	Home	Raina, Rogers, & Holm (2007)	
Sweep	Home		Raina, Rogers, & Holm (2007)
Indoor walking	Home		Raina, Rogers, & Holm (2007)
Home safety	Home	Raina, Rogers, & Holm (2007)	
Play bingo	Home $		Raina, Rogers, & Holm (2007)
Oven use	Clinic ¥		Raina, Rogers, & Holm (2007)
Stovetop use	Clinic		Raina, Rogers, & Holm (2007)
Use sharp utensils	Home		Raina, Rogers, & Holm (2007)
Clean up after meal preparation	Home ¥	Raina, Rogers, & Holm (2007)	

Note. FM = functional mobility; PC = personal care activities of daily living; CIADL = cognitive-oriented instrumental activities of daily living; PIADL = physical-oriented instrumental activities of daily living; * = comparisons at the task level was performed with Raina et al. only as it is the only study that used the PASS and compared performance between clinic and home at all three levels; ¥ = statistically significant; $ = substantially relevant but not statistically significant (DIF greater than 0.5 logits but $p > .05$).

4.0 DIFFERENCES IN TASK PERFORMANCE OVER TIME FOR OLDER WOMEN WITH MAJOR DEPRESSION

4.1 BACKGROUND AND SIGNIFICANCE OF THE PROBLEM

Old age is often associated with considerable debilitating cognitive and physical impairments (Turner & Noh, 1988). Over time, these impairments become more prevalent, severe, and complicated, thereby increasing vulnerability to affective disorders such as major depression (Beekman, Deeg, Braam, Smit, & Tilburg 1997). National data indicate that 1% to 5% of community-dwelling older adults have clinical depression (National Institute of Mental Health [NIMH], 2007). Of these, more than 40% have considerable difficulties in the performance of the activities of daily living (ADL) and instrumental activities of daily living [IADL] (Simon et al., 1998). These statistics indicate that the coexistence of age-related and affective impairments (i.e., clinical depressive symptoms) may amplify the difficulties in ADL and IADL (Katz, 1996). In fact, there is increasing evidence that older adults diagnosed with depression exhibit difficulties in daily tasks (Simon et al., 1998; Wells et al., 1989) considerably more than older adults without depression (Callahan et al., 1994), and that their difficulties are comparable to or worse than older adults with chronic medical illnesses (Wells & Burman, 1992). These difficulties also seem to persist over time, increasing the risk for loss of independence in ADL and IADL (Callahan et al., 1994; 1998).

Additionally, national statistics consistently show that women have higher prevalence rates of depression (NIMH, 2007) and loss of independence in daily tasks (Guralnik, Leveille, Hirsch, Ferrucci, & Fried, 1997) than men. Moreover, the percentage of women aged 65 years or older is projected to increase from 7 % in 1940 to 22.3 % in 2040: a tenfold increase (Guralnik et al., 1997). Therefore, given that an unprecedented number of women will live to a very old age in the coming century, and that many of them will exhibit depressive disorders and difficulties in task performance (Guralnik et al.) it is important to understand not only the global disability experienced by older women with depression but also which specific ADL and IADL are affected. This identification will help target areas for intervention, and improve the quality of life for this population. Hence, in this study we aim to examine changes (deteriorations, improvements) or stability in ADL and IADL task performance over 6 months in older women with depression. Deteriorations, improvement, or stability in task performance over time will shed the light on tasks that need follow-up assessments over time.

4.2 METHODS

4.2.1 Participants

We used a dataset collected in a previous study (AG08947) that recruited five diagnostic groups of older adults from the Benedum Geriatric Center, in Pittsburgh, Pennsylvania. Further details about the subjects in the previous study are provided in Chapter 3 of this Dissertation. For this study we used the data subset of the depression group (n = 56). The subjects in this data subset were older women with major depression who were assessed on their functional performance at home at the beginning of the study (Time-1) and six months later (Time-2).

4.2.2 Instruments

Task performance for the 56 older women was assessed with the Performance Assessment of Self-Care Skills [PASS] (Rogers & Holm, 1989). The PASS is a performance-based, criterion-referenced assessment tool which has 26 tasks classified into four functional domains: functional mobility (FM) with 5 items, personal care activities of daily living (PC) with 3 tasks, cognitive-oriented instrumental activities of daily living (CIADL) with 14 tasks, and physical-oriented instrumental activities of daily living (PIADL) with 4 tasks (see Table 4-1). Each task contains a group of subtasks with a total number of 163 subtasks for all items in the PASS. The PASS rates task performance for three constructs: independence (assistance needed while performing the task), safety (risks observed while performing a task), and adequacy (quality of final task outcome). However, in this study we only used the data on the independence construct. For each task in the PASS the assessor evaluates the independence in task performance based on the number of cues provided. The cues range from least to most assistive (see Table 4-2). The tasks are therefore rated on the frequency and level of cues provided by the assessor. Each subtask is given a score from 0 (total assistance) to 3 (no assistance/independent) (see Table 4-3). The overall PASS independence score is the mean score of all tasks, and the mean score for each task is the mean score of all subtasks for that tasks. The PASS was reported to be valid and reliable in several studies (see Chapter 3) for several older adults populations.

Table 4-1 PASS domains (4) and tasks (26)

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

Note. From “Disability in Older Adults with Depression” by D. Chisholm, 2005, doctoral dissertation. Copyright 2005 by Denise Chisholm. Adapted with permission.

Table 4-2 PASS prompts hierarchy

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

Note. From "Disability in Older Adults with Depression" by D. Chisholm, 2005, doctoral dissertation. Copyright 2005 by Denise Chisholm. Adapted with permission.

Table 4-3 PASS independence scoring criteria

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

Note. From "Disability in Older Adults with Depression" by D. Chisholm, 2005, doctoral dissertation. Copyright 2005 by Denise Chisholm. Adapted with permission.

4.3 DATA ANALYSIS

4.3.1 Participants

Data from all research subjects with major depression (n = 444) who were tested with the PASS were included in the item anchor dataset (AG08947; MH113679). The smaller dataset for Time-1 - Time-2 analysis is consisted of an intact sample of older women with major depression (n = 56) which had remitted, but they continued to be seen for medication management. The 56 older women were assessed on their daily performance at the beginning of the study (Time-1) and 6 months later (Time-2). Descriptive statistics of the demographic variables were calculated for the 56 subjects (see Chapter 3 for descriptive statistics for the item anchor dataset). We used SPSS

version 15.0 (SPSS Inc., 2002) to calculate all descriptive statistics for the demographic and impairment variables for both datasets.

4.3.2 Data preparation for Rasch analysis

Similar to Chapter 3 we used the large dataset to examine the construct validity and reliability of the PASS. We also used the Rasch model to prepare our data for statistical analysis, and WINSTEPS (Linacre, 2006) to perform all Rasch analyses for this study too. For consistency, we used the following order in explaining these analyses throughout the METHODS and RESULTS sections: data preparation (item difficulty estimate, item anchoring, person ability estimate, and item and person hierarchies); construct validity (fit statistics, principal component analysis and category function analysis); and reliability (item reliability index, person reliability index, person separation index, and internal consistency).

4.3.3 Data preparation: Item difficulty estimates

Using WINSTEPS we transformed the PASS ordinal scores into interval scores. The transformed scores are called the item difficulty logits (calculated from the total number of persons who fully performed a PASS item without assistance). As a result, each item in the PASS was given a logit which determined the relative difficulty of an item compared to the rest of the PASS items.

4.3.4 Data preparation: Item anchoring

After calculating a logit for the PASS items, WINSTEPS anchors (fixes) the difficulty estimate (logit) of these items to ensure item invariance or item calibration. Item invariance ensures a consistent item difficulty estimate of the PASS items for other samples similar in traits and size. This allowed for a more valid comparison across different samples (males and females or times

(Time-1 and Time-2). Anchored item difficulty logits were also used to examine the PASS construct validity and reliability.

4.3.5 Data preparation: Person ability estimate

Based on the anchoring item difficulty estimate, a person ability estimate in logits was determined for each of the 56 subjects. The person ability estimate is calculated from the total number of items that a person can perform independently (score 3 in the PASS). This ability estimate indicated the relative ability of a person compared to the other persons in the sample.

4.3.6 Data preparation: Item and person hierarchies

After calculating item difficulty and person ability estimates, WINSTEPS created two hierarchies: the item difficulty hierarchy (ordering items from easiest to most difficult) and the person ability hierarchy (ordering persons from least independent to most independent). The midpoint for both hierarchies was zero. According to the item hierarchy, items with greater logits are more difficulty than items with lower logits for that sample (and similar samples). According to the person hierarchy, persons with higher logits are more independent than persons with lower logits for that set of items.

4.3.7 PASS construct validity: Fit statistics

To examine the PASS construct validity we used WINSTEPS to calculate fit statistics which determine the compatibility of the PASS items, and persons in the sample, with the Rasch model expected values for items and persons. Fit statistics provide infit (inlier sensitive) and outfit (outlier sensitive) values as indicators of unidimensionality. Items (and persons) with infit and outfit values between 0.5 and 1.7 are considered compatible with the Rasch model. Items or

persons with infit or outfit values not within the acceptable range are referred to as noise, and are to be reinvestigated. For more details about fit statistics see Chapter 3.

4.3.8 PASS construct validity: Principal component analysis (PCA) of Rasch residuals

A principal component analysis (PCA) of Rasch residuals was also performed to further examine the unidimensionality of the PASS as a measure of construct validity. The PCA detects secondary dimensions in the data by calculating the units of variance that were explained by the Rasch model. It also calculates the unexplained variance of the first contrast (or the largest secondary dimension identified by PCA). The more variance explained by the model (should be larger than 60%) and the less unexplained variance in the first contrast (should be less than 3.0 units or less than 5% of the total variance), the more unidimensional the construct under investigation.

4.3.9 PASS construct validity: Category function analysis

Category function analysis was also used to examine the construct validity of the PASS. This analysis examined the appropriateness (functionality) of the 4 response categories in the PASS (0, 1, 2, and 3). There are four features to be considered a category function analysis: 1) category frequency, which is the number of responses for each category; a minimum frequency of 10 response per each category is recommended, 2) average measures, which is the average person ability logits for persons who scored a specific category; a monotonic increase in the category is recommended, 3) thresholds, which is the task difficulty estimate for choosing one category over another (should also be monotonic), and 4) category fit, which are infit and outfit statistics and both should be less than 2.

4.3.10 PASS reliability: Item reliability index

The item reliability index represents future replicability of the item ordering if these items were assessed in another sample with similar traits and size. Higher item reliabilities indicate greater stability of item ordering.

4.3.11 PASS reliability: Person reliability index

The person reliability index also represents future person ability ordering if the same persons were assessed on a similar set of items that measured the same construct. Higher person reliabilities indicate greater stability of person ordering (a value $\geq$ .80 indicates greater stability of person ordering).

4.3.12 PASS reliability: Person separation index

The person separation index is a measure of the stability of person ordering. A higher separation index indicates more reliable differentiation between persons of higher ability and those of lower ability. Values close to or higher than 2, represent a good person separation index.

4.3.13 PASS reliability: Internal consistency

Cronbach's alpha indicates the internal consistency of the PASS. Values > 0.90 indicate high internal consistency. Cronbach's alpha as well as person separation reliability helped in investigating the reliability of the PASS independence measure.

4.3.14 Performance at Time-1 versus Time-2 at the PASS overall level

After preparing our data we examined the differences between the overall task performance at Time-1 and that at Time-2 using paired (or dependent) t-tests. Person ability estimates were used

for this analysis to help understand the differences in task performance over time for older women with major depression. Moreover, as older adults with major depression usually have somatic and cognitive impairments that coexist with the affective impairments, and may eventually add to the influence of time, we decided to investigate the correlation between change in person ability logits from Time-1 to Time-2 and change in six potential confounding variables. Significant correlations may indicate a confounding effect of these six variables on the task performance over time. In contrast, non significant correlations may indicate that change in task performance over time is mainly counted for by time. The five factors (six measures) were age, change in physical function (change in total score of the Keitel Functional Test [KFT]), change in cognitive function (change in total score of the Trail Making Test-A [TMT-A], Trail Making Test-B [TMT-B], and the Modified Mini-Mental State Examination [3MS]), and depressive symptoms (change in total score of the Geriatric Depression Scale [GDS]). The KFT consists of 24 items that assess physical limitations (restricted range of motion and strength) in the upper and lower extremities, and the spinal column with a total score ranging from 4 [minimal limitations] to 100 [total disability] (Eberl, Fasching, Rahlfs, Scheleyer & Wolf, 1976). The TMT-A requires a subject to draw a line to connect 25 consecutive numbers (1 to 25) and measures visual scanning, number recognition, numeric sequencing, and motor speed (Giovagnoli et al., 1996). TMT-B requires a subject to draw a line that connects numbers and letters in sequence (e.g., 1-A, 2-B), and measures complex attention and shifting ability (Arbuthnott & Frank, 2000; Giovagnoli et al., 1996). The time in seconds needed to complete TMT-A and TMT-B is the subject's score for that test. Subjects with scores lower than the 50^{th} percentile for their age and education level were considered slower performers, and those with scores higher or equal to the 50^{th} percentile for their age and education level were considered fast performers (Tombaugh, 2004). The GDS consists of 15 questions about

the presence of depressive symptoms with higher scores (≥ 5) indicating more depressive symptoms (Yesavage et al., 1982). The 3MS consists of 100 questions that assess cognitive functions such as orientation to time and place, naming ability, and recall, with a final score ranging from 30 and 100 (higher score indicates better cognitive functioning [McDowell, Kristjunsson, Hill & Hébert, 1997; Teng & Chui, 1987]). We conducted paired t-tests for each confounding factor at Time-1 and Time-2 to identify changes in these factors that may contribute to the change in task performance over time. Furthermore, we conducted multiple regression analysis to consider the additive effect of all the confounding factors combined. The results of the regression will show the factors that mainly contributed to the variance in the person ability for task performance over time.

4.3.15 Performance at Time-1 versus Time-2 at the PASS domain level

We used the differential group functioning (DGF) framework to examine the differences between the person abilities of the two groups across groups of items. For our study, the groups are persons at Time-1 and Time-2, and the groups of items are the four PASS domains (FM, PC, CIADL, and PIADL). We conducted four paired t-tests (with Bonferroni corrections) to compare differences in task performance for the four domains at Time-1 and Time-2. These four analyses helped us to understand significant differences in person ability, by domain, over time.

4.3.16 Performance Time-1 versus Time-2 at the PASS task level

We also explored differences in person ability over time at the task level which helped us identify specific tasks that are more likely to improve or deteriorate over time for older women with major depression. As there are 26 items in the PASS we needed a powerful analysis to detect differences at this level. Therefore, we decided to use a method provided by the Rasch model called

Differential Item Functioning (DIF). The DIF examines the invariance of item difficulty estimate by determining significant differences between two groups of people (Time-1 versus Time-2). Hence, for this analysis, we used the item difficulty estimate instead of person ability estimate to identify the differences in the perceived item difficulty over time. In other words, DIF identified which items people found significantly different at Time-1 from at Time-2. The difference between item difficulty estimate at Time-1 and that at Time-2 is known as DIF contrast. Only DIF contrasts of 0.50 logit or higher, with a significance value of $p < 0.05$ are considered significant differences between two groups.

4.4 RESULTS

4.4.1 Participants

The demographic variables for the large depression dataset (n = 444) were reported in Chapter 3 (see Table 3-4). Demographic variables for the small data subset of Time-1 and Time-2 (n = 56) are summarized in Table 4-4. All subjects were females with a mean age 75.73 years. The demographic variables showed that the majority of subjects were Caucasian (81.4%), widowed (62.7%), and lived alone (62.7%). Additionally, most subjects had a high school education or higher (81.3%), and had a household income $\geq$ \$10,000 (62.7%). The impairment variables showed that physical impairment was low (KFT), performance speed was moderate (TMT-A, TMT-B), depressive symptoms were mild (GDS), and dementia was not evident ($\geq$78 on 3MS). Subjects were diagnosed with major depression, which had remitted at Time-1; however, they continued to be treated for medication management at Time-2 in an outpatient clinic.

Table 4-4 Descriptive statistics for Time-1 and Time-2 data subset

Variables	*n* = 56	
Demographic variables		
Age (years)	75.82	
Ethnicity (% Caucasian)	81.40	
Marital status (% widowed)	62.70	
Living status (% living alone)	62.70	
Education (% ≥ high school)	81.30	
Household income (% ≥ $10,000)	62.70	
Impairment variables*	Time-1 Mean (SD)	Time-2 Mean (SD)
KFT (score range 4 to 100)	21.63 (8.68)	21.73 (9.05)
TMT-A (scores in seconds)	52.56 (17.85)	54.78 (25.83)
TMT-B (scores in seconds)	125.80 (55.23)	113.64 (52.70)
GDS (score range 0 to 15)	3.58 (3.33)	5.65 (1.55)
3MS (scores range 0 to 100)	92.93 (5.13)	90.77 (7.24)

Note. SD = standard deviation; KFT = Keitel Functional Test; TMT-A = Trail Making Test-A; TMT-B = Trail Making Test-B; GDS = Geriatric Depression Scale; 3MS = Modified Mini-Mental State Examination; * = higher scores in KFT, TMT-A, TMT-B, and GDS indicate greater impairment, and lower scores in 3MS indicate less impairment.

4.4.2 Data preparation: Item difficulty estimates

All PASS item raw scores were transformed into logits that represented the item difficult estimates which were then were used to create and anchor the item difficulty hierarchy (see Table 3-6 in Chapter 3).

4.4.3 Data preparation: Item anchoring

The item hierarchy was anchored using the larger depression dataset (n = 444). Among the 26 PASS items trimming toenails was the most difficult and indoor walking was the easiest (See Table 3-6 in Chapter 3).

4.4.4 Data preparation: Person ability estimate

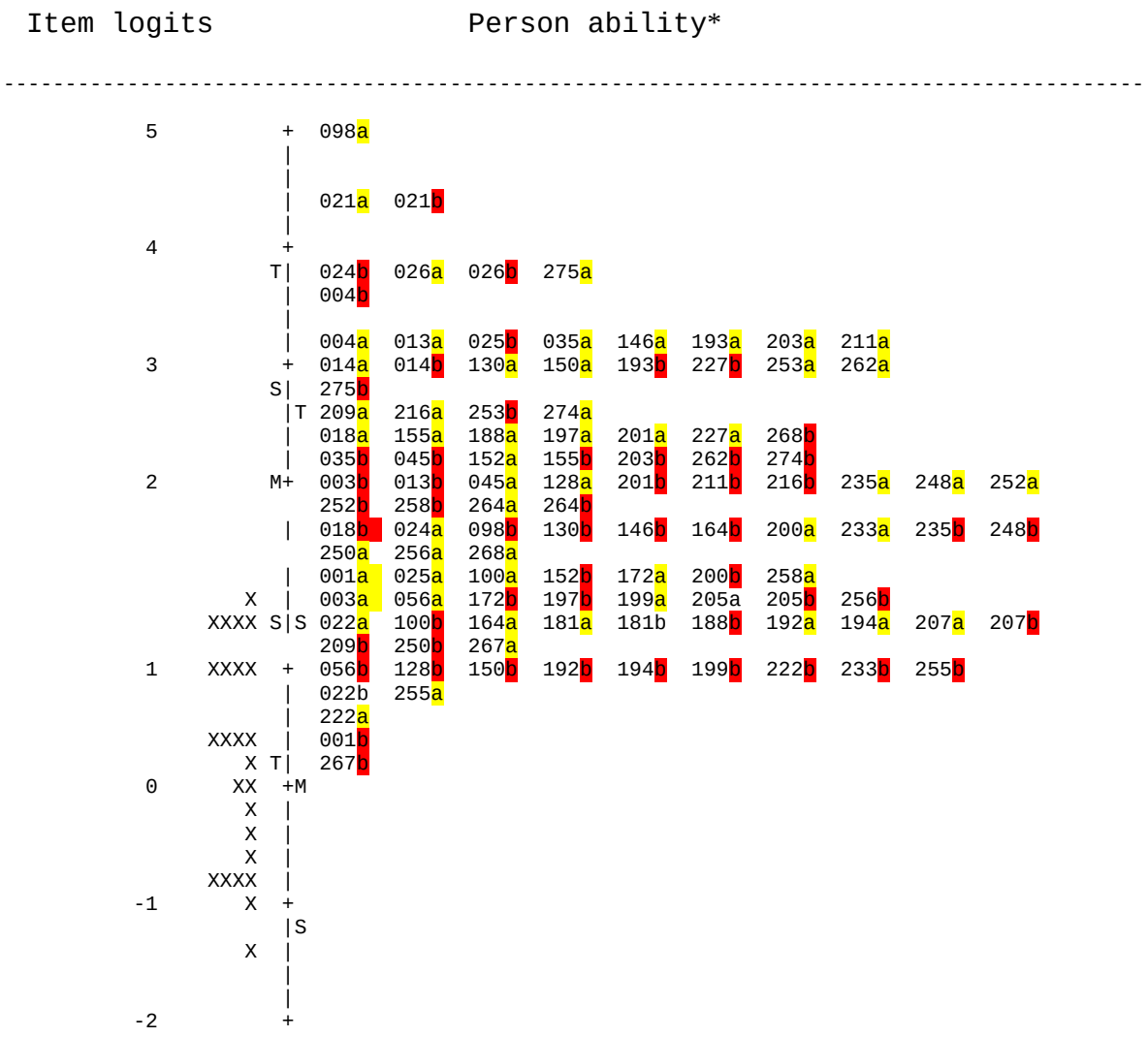
Person ability was calculated for all 56 persons at Time-1 and Time-2. Figure 4-1 illustrates the distribution of person ability at Time-1 and Time-2. Figure 4-2 represents the item-person map which places each person in the sample (based on the person ability estimate) against the anchored hierarchy of item difficulty.

4.4.5 Data preparation: Item and person hierarchies

Both item and person hierarchies were created after anchoring item difficulty estimates (See Table 3-6 in Chapter 3). This allowed for a valid comparison between the performance at Time-1 and Time-2 for the 56 subjects (Figure 4-1 and Figure 4-2).

Best Performance (Most Independent)		
Person ability logit at Time-1 (n = 56)		Person ability logit at Time-2 (n = 56)
5.71	5.00	
4.47	4.00	4.47
	3.75	
3.73; 3.73		3.70; 3.73; 3.73
	3.50	
3.27; 3.27; 3.27; 3.27	3.25	3.27
3.24		
3.19; 3.19	3.00	
2.94; 2.94; 2.94; 2.94; 2.94		2.94; 2.94; 2.94
	2.75	2.82
2.67; 2.67; 2.59	2.50	2.67
2.45; 2.45; 2.45; 2.45; 2.42; 2.37		2.45
	2.25	2.26; 2.26; 2.26
2.10	2.00	2.10; 2.10; 2.10
1.96;1.96; 1.96; 1.96; 1.96; 1.92 1.83; 1.83; 1.83		1.83; 1.83;1.83; 1.83;1.89; 1.96; 1.96; 1.96; 1.96; 1.96; 1.96; 1.96; 1.96
	1.75	
1.71; 1.71; 1.71; 1.68; 1.61; 1.61;		1.71; 1.71
1.52; 1.51	1.50	1.52; 1.52
1.43; 1.43		1.43; 1.43; 1.48
1.35; 1.30; 1.27; 1.27;1.27	1.25	1.27;1.35
		1.20; 1.20; 1.20; 1.20
1.13; 1.13; 1.13; 1.13		1.13
	1.00	1.00; 1.00; 1.01; 1.06; 1.06; 1.06; 1.06; 1.08
0.83		0.83; 0.94
	1.75	
0.57	0.50	
	0.25	0.47
		0.22
	0.00	
Worst Performance (Least Independent)		

Figure 4-1 Person ability estimates (in logits) for persons at Time-1 and Time-2



Note. Refer to Table 3-6 in Chapter 3 for item difficulty hierarchy;* Yellow= person ability at Time-1; Red= person ability at Time-2.

Figure 4-2 Item-person map. The map illustrates the person ability (right side) for persons at Time-1 (a) and Time-2 (b) against the anchored item difficulty hierarchy (left side)

4.4.6 PASS construct validity: Fit statistics

Most of the PASS items were within the acceptable range (0.5 to 1.7). Only trimming toenails and stair use slightly deviated from the acceptable range. This indicates that most of the PASS items are compatible with the Rasch model (see Table 3-6 in Chapter 3).

4.4.7 PASS construct validity: Principal component analysis (PCA) of Rasch residuals

The PCA clearly illustrated the unidimensionality of the PASS construct of independence. The Rasch model explained 84.9 % of the total variance (100%), and the first contrast explained only 1.4% (considerably less than 5% which considered noise) of the unexplained variance (see Figure 3-3 and Figure 3-4 in Chapter 3).

4.4.8 PASS construct validity: Category function analysis

All features in the category function analysis met the recommended criteria for appropriate categories functioning (see Table 3-8 and Figure 3-5 in Chapter 3). That is, each PASS category had more than 10 responses, higher categories corresponded with higher person abilities, and fit statistics for all categories were less than 2 logits.

4.4.9 PASS reliability: Item reliability index

The calculated item reliability index was 0.98 which indicated excellent reliability of the PASS in predicting the item ordering if this group of items were used with other samples of similar traits and size.

4.4.10 PASS reliability: Person reliability index

The person reliability index was calculated as 0.76 which indicated the stability of the PASS in predicting future person ordering, if a similar set of items (measuring the same construct) were assessed for the sample.

4.4.11 PASS reliability: Person separation index

The person separation index was 1.76 which indicated the ability of the PASS to reliably differentiate between persons with higher abilities from those with lower abilities.

4.4.12 PASS reliability: Internal consistency

The PASS showed excellent internal consistency with Cronbach's alpha of 0.90.

4.4.13 Performance at Time-1 versus Time-2 at the PASS overall level

Table 4-5 shows the *t*-test results of the differences in the PASS overall between Time-1 and Time-2. The person ability logits indicated that persons were significantly more independent at Time-1 than at Time-2 for overall PASS performance ($p < .002$). Moreover, none of the correlations between the change score in person ability logits from Time-1 to Time-2 and age or the change scores in the five potential confounding factors were significant. Hence, changes in task performance over time was best accounted for by time (see Table 4-6).

Table 4-5 Comparison between Time-1 and Time-2 overall person ability PASS scores

Time	Mean*	*t*	*df*	*SE*	Significance
Time-1	2.23	3.27	55	0.12	$p < .002$
Time-2	1.85				

Note. *df* = degrees of freedom; *SE* = standard error; * = mean person ability logit scores for each group.

Table 4-6 Correlations between change in person ability from Time-1 and Time-2 and age and change in five confounding factors

	Age	KFT	TMT-A	TMT-B	GDS	3MS
Change in person ability	r = -.13 ($p < .17$)	r = -.16 ($p < .12$)	r = .07 ($p < .30$)	r = .04 ($p < .41$)	r = .03 ($p < .42$)	r = -.04 ($p < .39$)

Note. r = Pearson correlation; KFT = Keitel Functional Test; TMT-A = Trail Making Test-A; TMT-B = Trail Making Test-B; GDS = Geriatric Depression Scale; 3MS = Modified Mini Mental State Examination.

Table 4-7 summarizes the t-test results for five confounding factors. The results showed that of the five factors, only depressive symptoms significantly increased over time ($p < .001$ with Bonferroni corrections). Hence, the subjects had remitted depressive symptoms at Time-1 (GDS < 5), but showed mild depressive symptoms at Time-2 (GDS > 5). Similarly, impairments in physical (KFT scores) and cognitive (TMT-A and 3MS) abilities also increased slightly over time, but did not reach statistical significance. TMT-B scores were the only factor that showed slight improvement over time, but they also did not reach statistical significance.

Table 4-7 Comparison of five confounding factors between Time-1 and Time-2

Time	Mean	*t*	*df*	*SE*	Significance
KFT-T1 KFT-T2	21.63 21.73	-0.17	55	0.64	$p < .87$
TMT-A-T1 TMT-A-T2	52.56 54.78	-0.83	54	2.67	$p < .41$
TMT-B-T1 TMT-B-T2	125.80 113.64	1.60	43	7.62	$p < .12$
GDS-T1 GDS-T2	3.58 5.65	-5.00	54	0.41	$p < .001$*
3MS-T1 3MS-T2	92.93 90.77	2.36	55	0.91	$p < .02$**

Note. T1 = Time-1; T2 = Time2; SE = standard error; KFT = Keitel Functional Test; TMT-A = Trail Making Test-A; TMT-B = Trail Making Test-B; GDS = Geriatric Depression Scale; 3MS = Modified Mini-Mental State Examination; * = significant at p < .01 with Bonferroni corrections; ** = not significant after Bonferroni corrections.

Table 4-8 summarizes the results of the multiple regression analysis. The results showed that the effect of the confounding variables combined did not significantly contribute (R square = 0.06; F = 0.39; p < 0.88) to the explanation of the variance in person ability over time. Moreover, the results showed a negative adjusted R square which indicated that the confounding variables are considered “noise” or distracters.

Table 4-8 Results of the multiple regression analysis

Model	R square	Adjusted R square	Sum of Squares	df	Mean square	F	Sig
Regression	0.06	-.095	2.14	6	0.36	0.39	0.88
Residual	_	_	32.88	36	0.91	_	_
Total	_	_	35.02	42	_		

Note. df = degrees of freedom.

4.4.14 Performance at Time-1 versus Time-2 at the PASS domain level

Table 4-9 summarizes the differences in person ability across the four domains at Time-1 and Time-2. Performance was significantly more independent at Time-1 than at Time-2 for CIADL ($p < .005$) and PIADL domains ($p < .001$) after Bonferroni corrections. Moreover, performance in PC was more independent at Time-1 than Time-2 although the results did not reach statistical significance. Differences between Time-1 and Time-2 for FM were not significant.

Table 4-9 Difference between Time-1 and Time-2 PASS person ability domain scores

PASS domains	Environment	Mean*	*t*	*df*	Significance **
FM	Time-1	1.58	-0.19	55	p <0.85
	Time-2	1.61			
PC	Time-1	1.78	1.65	55	p <0.11
	Time-2	1.48			
CIADL	Time-1	2.94	2.93	55	p <0.005
	Time-2	2.54			
PIADL	Time-1	2.46	3.57	55	p <0.001
	Time-2	1.91			

Note. FM = functional mobility; PC = personal care activities of daily living; CIADL = cognitive-oriented instrumental activities of daily living; PIADL = physical -oriented instrumental activities of daily living; *df*= degrees of freedom; * = person ability logits; ** significant at $p < 0.013$ with Bonferroni corrections.

4.4.15 Performance at Time-1 versus Time-2 at the PASS task level

Table 4-10 summarizes the DIF results. Performance was significantly more independent at Time-1 than at Time-2 for stovetop (CIADL) use and clean up after meal preparation (PIADL), whereas performance was significantly more independent at Time-2 than at Time-1 for bathtub/shower transfer (FM) and home safety (CIADL). Moreover, performance was

substantially relevant (more independent with a DIF contrast greater than 0.50 logits), but not statistically significant ($p > 0.05$) at Time-1 for FM items stair use and indoor walking, and PIADL items bending, lifting, carrying and sweeping. Finally, performance was substantially relevant (more independent) but not statistically significant at Time-2 for toilet transfer (FM) and playing bingo (CIADL). Performance on the remaining PASS tasks remained the same with DIF contrasts equal to zero or less than 0.5 logits. Figure 4-3 visually illustrates the DIF measures for Time-1 and Time-2. Figure 4-4 summarizes the results about change in task performance over time at the overall, domain, and task levels.

Table 4-10 Differences in person ability at Time-1 and Time-2 at the task level

PASS tasks	DIF contrast	Probability	Less difficult in
Bed transfer	0.47	p < 0.31	Time-2
Stair use	0.54	p < 0.12	Time-1
Toilet transfer	0.86	p < 0.16	Time-2
Oral hygiene	0.07	p < 0.87	Time-1
Bathtub/shower transfer	0.58	p < 0.02*	Time-2
Trim toenails	0.00	p < 0.99	Time-1= Time-2
Dress	0.43	p < 0.34	Time-1
Shop	0.19	p < 0.42	Time-1
Pay bills by checks	0.29	p < 0.25	Time-2
Balance a checkbook	0.03	p < 0.89	Time-2
Mail bills and checks	0.16	p < 0.56	Time-1
Bending, lifting, carrying	0.71	p < 0.13	Time-1
Telephone use	0.12	p < 0.85	Time-2
Medication management	0.33	p < 0.22	Time-2
Change bed linens	0.19	p < 0.46	Time-1
Obtaining information: auditory	0.30	p < 0.65	Time-1
Obtaining information: visual	0.04	p < 0.94	Time-2
Small repairs	0.20	p < 0.49	Time-2
Sweep	0.71	p < 0.26	Time-1
Indoor walking	1.54	p < 0.57	Time-1
Home safety	0.76	p < 0.02*	Time-2
Play bingo	1.47	p < 0.21	Time-2
Oven use	0.33	p < 0.24	Time-2
Stovetop use	1.26	p < 0.001*	Time-1
Use sharp utensils	0.17	p < 0.64	Time-2
Clean up after meal preparation	1.88	p < 0.001*	Time-1

Note. DIF= differential item functioning; * = item difficulty estimate is statistically significantly different between Time-1 and Time-2 for that item.

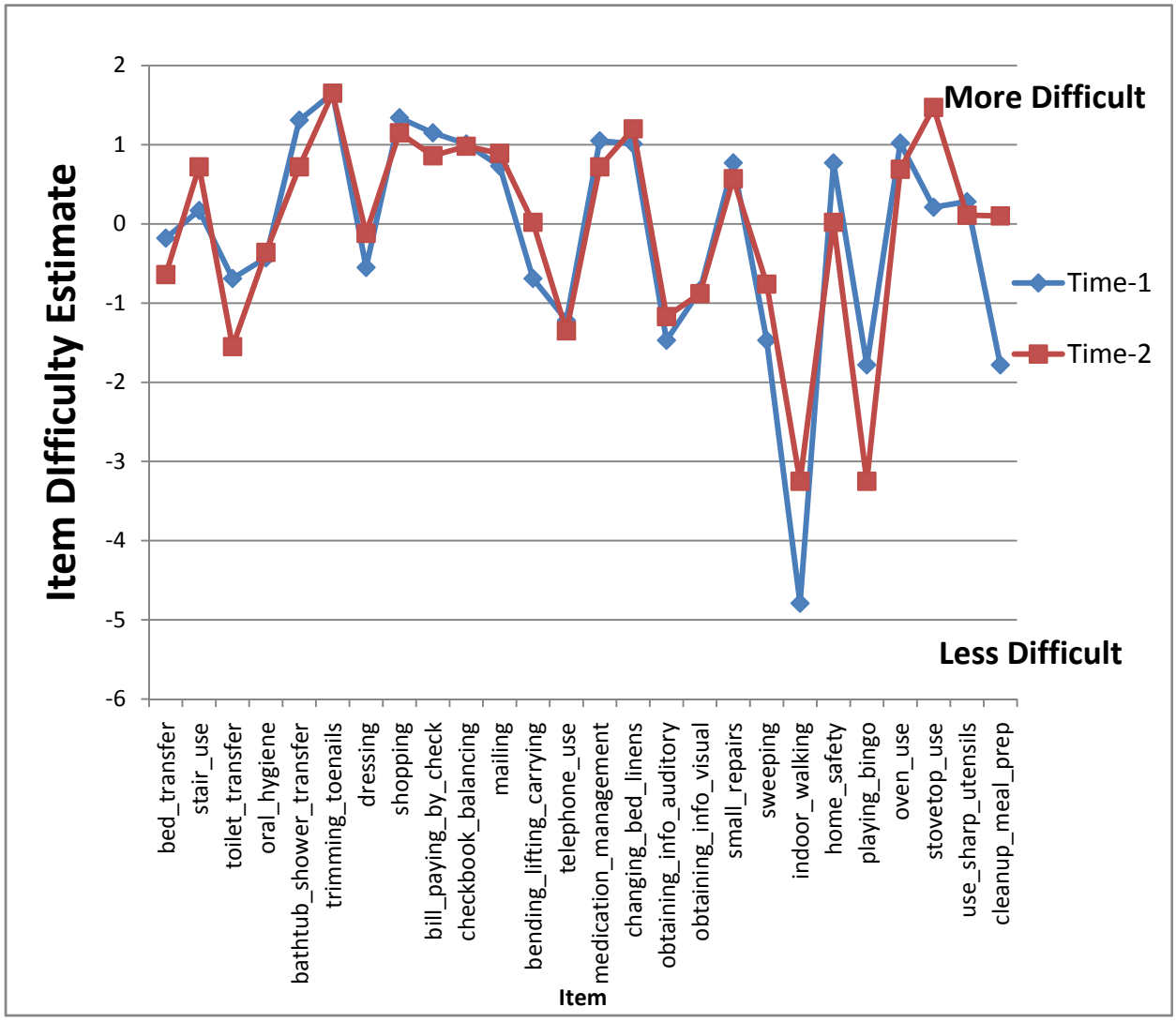
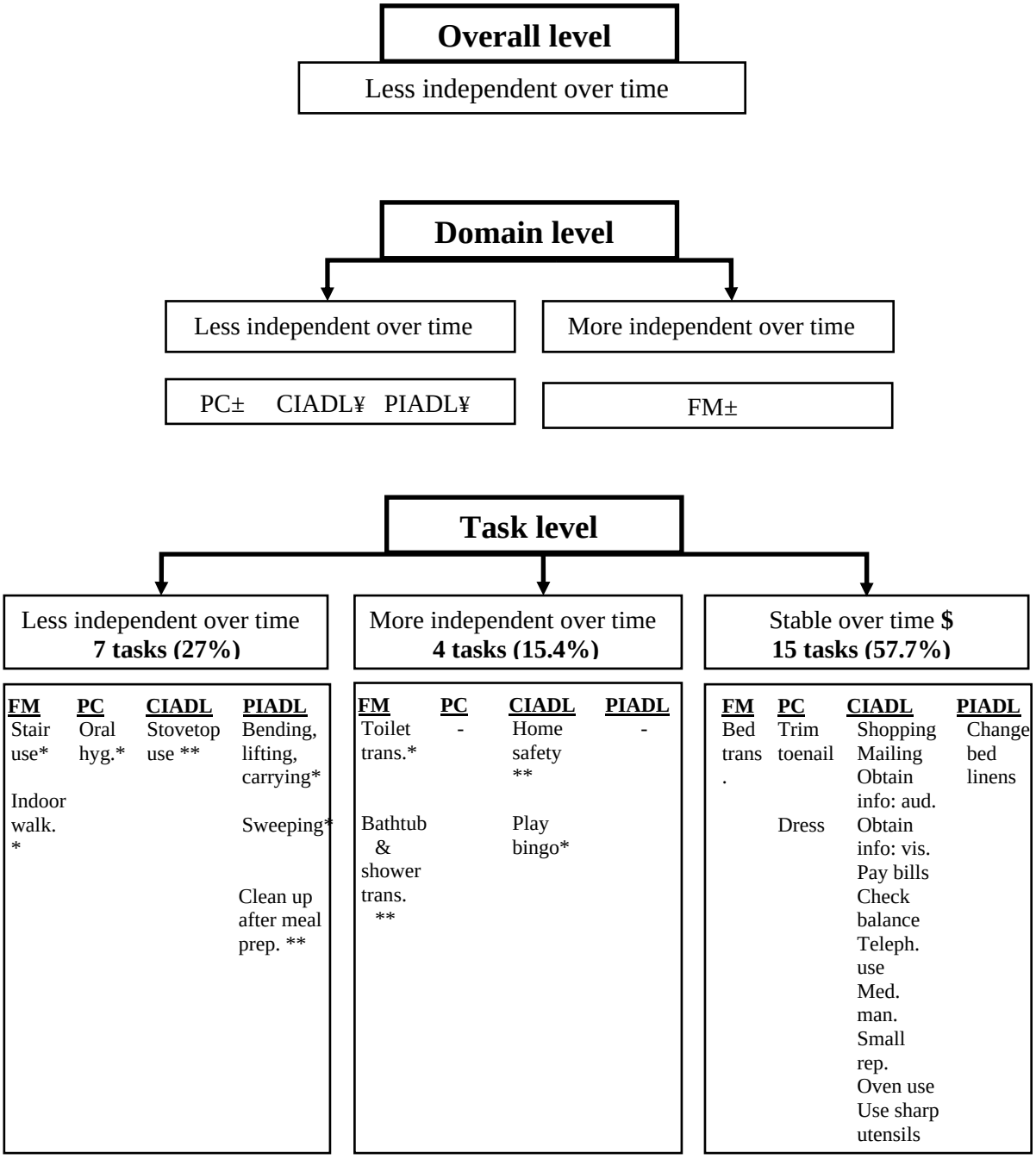


Figure 4-3 Comparison between the item difficulty estimates at Time-1 and Time-2



Note. FM = functional mobility; PC = personal care activities of daily living; CIADL = cognitive-oriented instrumental activities of daily living; PIADL = physical-oriented instrumental activities of daily living; ± = not significant; ¥ = statistically significant ($p < .013$ with Bonferroni corrections); * = substantially relevant (DIF contrast greater than 0.50) but not statistically significant ($p > .05$); ** = statistically ($p < .05$) and substantially relevant (DIF contrast greater than 0.50); $ = DIF equals to zero or less than 0.50 logits and therefore no substantive change on task performance can be more or less independent at Time-1 or Time-2 as shown in table 4-6. Refer to table 4-6 for tasks' full names.

Figure 4-4 Summary of results about change in task performance over time at the overall, domain, and task levels

4.5 DISCUSSION

Older adults comprise a rapidly growing segment of the population. A large portion of that segment exhibit several limitations in ADL and IADL, and these limitations seem to increase over time. Our findings showed that independence in performance of ADL and IADL, tended to deteriorate over a short time period of 6 months for older women with remitted symptoms of major depression. In this study, our aim was first to investigate the effect of time on the overall performance of ADL and IADL. Then, we aimed to identify functional domains and individual tasks that may be most susceptible to deterioration over time for older women with depression.

Our study was the first to use a sample of older women with major depression to examine changes over time in task performance at three levels of performance: overall, domain, and task levels. Previous studies were longitudinal (Hays, Steffens, Flint, Bosworth, & George, 2001; Rogers, Holm, Goldstein, McCue, & Nussbaum, 1994), epidemiological (Bowling, Farquhar, & Grundy, 1996; Chou, 2005; Gallo, Rabins, Lyketsos, Tien, & Anthony, 1997; Gallo, et al., 2003; Heikkinen & Kauppinen, 2004; Kennedy, Kelman, & Thomas, 1990; Laukkanen et al., 1994; Penninx, Leveille, Ferrucci, van Eijk, & Guralnik, 1998; Prince, Harwood, Thomas, & Mann, 1998; Williamson & Schulz, 1992; Zeiss, Lewinsohn, Rohde, & Seeley, 1996), or cross-sectional (Laukkanen, Heikkinen, Schroll, & Kauppinen, 1997; Mulsant, Ganguli, & Seaberg, 1997; Seaburn, Lyness, Eberly, & King, 2005; Steffens, Hays, & Krishnan, 1999) and used mixed

samples (males and females) of older adults with depressive symptoms rather than clinical depression (Bruce, Seeman, Merrill, & Blazer, 1994; Gill et al., 2004; Guccione et al., 1994; Han, 2002; Han & Jylha, 2006; Hays et al., 1997; Kurland et al., 2006; Lenze et al., 2001; Lenze et al., 2005; Ostir et al., 2001; Turner & Noh, 1988; Whyte et al., 2004). Even other relevant longitudinal studies that recruited older adults with clinical depression (Hays, Steffens, Flint, Bosworth, & George, 2001; Rogers, Holm, Goldstein, McCue, & Nussbaum, 1994) used mixed samples of males and females and assessed task performance at either the domain level (Rogers et al., 1994) or at the domain and task levels (Hays et al., 2001). Hence, no studies assessed task performance over time at the overall, domain, and task levels for a sample of older women with major depression.

In our study, performance at the three levels was assessed with the PASS; a performance-based, criterion-referenced, assessment tool. The PASS was reported to be valid and reliable in numerous studies (Chisholm, 2005; Finalyson, Havens, Holm, & Van Denend, 2003; Holm & Rogers, 1999; Holm, Rogers, & James, 2003; Rogers & Holm, 2000; Rogers, et al., 2003; Rogers, Holm, Goldstein, McCue, & Nussbaum, 1994; Rogers, Holm, Beach, Schulz, & Starz, 2001). However, in this study, because we used Rasch analysis, we confirmed the validity and the reliability of the PASS before completing further analyses. The PASS showed excellent validity, and reliability for assessing task performance in older women with major depression.

After verifying the appropriate use of the PASS for our sample we compared the person ability of our participants at two points of time (Time-1 and Time-2). This comparison allowed us to understand whether or not the person ability of older adults with major depression deteriorates over time, and if so, at which level of performance.

Generally, our findings at all three levels indicated that person ability deteriorated over just a 6-month time period in older women with depression.

Overall, performance was significantly less independent over time. Although remitted, depression still affected independence in daily activities even after considering age, physical impairment (KFT), and cognitive impairments (TMT-A, TMT-B, and 3MS) including those associated with executive functioning (TMT-B). Even after considering the additive effect of the confounding factors combined, time was still the main factor affecting task performance over 6 months. Hence, despite the fact that these women were becoming older and experiencing the gradual loss of physical and cognitive capacities (increases in KFT, TMT-A, and 3MS) that may have directly affected task performance over time, the burden of mild depressive symptoms over time was the main debilitating effect on task performance.

Second, our findings at the domain level showed that performance also deteriorated for several groups of items. That is, performance significantly deteriorated over time for CIADL and PIADL but not for FM (see Figure 4-4). The noticeable loss of independence over time in CIADL, and PIADL may possibly be explained by the fact that most cognitive IADL (e.g., paying bills) and physical IADL (e.g., sweeping) are complex tasks that require a combination of cognitive and physical abilities in order to be performed, as well as the motivation to carry them out. This combination of abilities is often compromised in older adults with depression due to the cognitive (difficulties in thinking, concentration, decision making, and memory) (Callahan, 2005; Turner & Noh, 1988), physical (fatigue, psychomotor retardation, restlessness, and sleep disturbances) (Turner & Noh, 1988), and affective symptoms (loss of motivation, apathy, hopelessness,

and persistent sadness) (American Psychiatric Association, 2000) experienced by this population. In contrast, independence in FM improved over time, although it did not reach statistical significance. In part, this finding also was not surprising as some tasks of FM (e.g., bed transfer) usually require "gross" physical motor abilities rather than a combination of several types of abilities. The daily use, per se, of these abilities in tasks such as bed mobility or transfer may be sufficient enough to maintain, or even improve, the performance of these tasks. However, some tasks of FM can be significantly challenging (e.g., stair use) and therefore improvement in such tasks may be surprising and not typically expected. In regard to pervious relevant studies, our findings were consistent with Rogers et al. (1994) in showing a significantly less independent performance over time in CIADL and PIADL. However, our results were not consistent with Hays et al. (2001) where most of the subjects showed improved performance over time in IADL domain, bearing in mind that they used self-report, not performance-based assessment.

Finally, at the task level the results varied with performance being less independent over time in 7 tasks, more independent over time in 4 tasks, and stable over time in 15 the tasks (see Figure 4-4). Less independence was noticed in most PIADL tasks followed by CIADL, FM, and PC. More independence over time was noticed in some CIADL and FM tasks; however, stable performance was noticed in several CIADL tasks. Although performance in most of the tasks did not dramatically change over time, the number of tasks in which performance was less independent has fundamental clinical significance. That is, less independence was noticed in stair use, indoor walking, oral hygiene, stovetop use, bending, lifting, and carrying, sweeping, and clean up after meal

preparation. This is not surprising as the common ability required for almost all these tasks is standing, tolerance for which tends to deteriorate in older adults in general and older adults with depression in particular (Penninx et al., 1998). Interestingly, however, most tasks in which performance was more independent or remained stable over time did not require, in a sense, "standing" to successfully complete the task (e.g., medication management). Hence, our findings suggest that tasks that require physical abilities (mainly standing), should receive extra attention and be followed over time for older women with depression. It is noteworthy, that although the results at the domain level showed that performance in the FM domain improved over time (not statistically significant), this improvement was primarily due to one task: bathtub/shower transfers.

Although none of the previous research samples matched ours, our findings were consistent with some of the findings in previous research, depending on the level of performance. No previous studies analyzed changes in performance, over time, at the overall level. At the domain level, our results were consistent with Rogers, Holm, Goldstein, McCue, and Nussbaum, (1994) which also showed significant loss of independence over time in CIADL and PIADL. However, unlike Rogers et al. (1994) our results did not show a significant loss of independence in FM or PC. Also at the domain level, our results were consistent with those of Hays, Steffens, Flint, Bosworth, and George (2001) who reported a non significant loss of independence in ADL (they combined PC and FM). In contrast to our findings, however, Hays et al. (2001) reported improvement in IADL over time. Finally, at the task level, our results were consistent with Hays et al. (2001) in showing stable performance over time for dressing, shopping, money management tasks, and meal preparation tasks. However, Hays et al. found toilet

and bathtub transfers to be stable, our sample improved over time. Similarly, Hays et al. found walking, grooming, and bending to be stable, whereas our sample decreased in independence over time.

As with all studies, this study had limitations. The use of a previously collected dataset limited certain factors such as the sample size, the time over which task performance was assessed, and the inclusion of other confounding variables that may have contributed to the change in performance (e.g., social support). In addition, although important and informative, the use of a sample of only women limited generalizing the results to men and women or of men only. Therefore, future studies are needed to confirm these results and examine task performance with larger samples and over longer times in older women with depression. Moreover, future studies that use samples of older men with major depression are needed to compare the effect of time between men and women. Finally, future studies that use more sophisticated impairment measures (e.g., Hamilton Depression Scale; an executive functioning battery of tests) are needed, especially if they are sensitive to change in people with depression.

In summary, our results showed that time can significantly negatively affect task performance at several levels in older women with major depression. This finding emphasizes the need for healthcare practitioners and functional assessment raters to not rely only on global measures of daily activities but also on specific task measures that include items that are known to deteriorate quickly over time in older women with depression. The identification of these tasks can help provide proper intervention to maximize independence for older women with major depression.

5.0 CONCLUSION

The purpose of this dissertation was to examine task performance in older women with major depression. The specific aims were to:

(1) Find and evaluate evidence on the influence of environment and the influence of time on task performance of older women with major depression.

(2) Examine the influence of two environments, the clinic and the home, on task performance of older women with major depression.

(3) Examine the effect of time on task performance of older women with major depression.

The first study (Chapter 2) had two objectives: a) to locate and evaluate evidence on the influence of two environments, the clinic and home, on task performance in older women with major depression, and b) to locate and evaluate evidence on the influence of time on task performance in older women with major depression. For the first objective, only limited relevant evidence was found. The research showed conflicting results regarding the evidence on the influence of the environment on task performance in various samples of older adults. That is, depending on the level of performance, some studies showed that the

clinic was more facilitative than the home for task performance, some showed that the home was more facilitative than the clinic, and some showed that the clinic and home had equal influence. Second, none of these studies used a sample of older adults with major depression despite the significant functional disability associated with major depression. All the studies used samples of older adults with physical conditions such as stroke, osteoarthritis, and heart failure. Furthermore, none of the studies used a sample of older women with major depression despite the higher prevalence of depression in women compared to men. All studies reviewed used mixed samples of males and females of older adults. Finally, most of these studies had methodological flaws that limited the generalization of their results to older adults with major depression. The use of different assessment methods in the clinic and home, poor definition of tasks assessed, and lack of demographic information are a few examples of these flaws. Hence, for these reasons, we conducted a study (Chapter 3) in order to examine the influence of environment on task performance in older women with major depression.

For the second objective, we also found limited research. Most of the studies found were cross-sectional or epidemiological and hence did not provide information on task performance over time. Moreover, most of the studies assessed global disability and did not specify functional domains and tasks, or assessed physical impairments rather than task-related functional limitations. Further, most of the studies used screening tools that detected depressive symptoms, rather than diagnostic tests that detected clinical depression. Even the few studies that used samples of older adults with clinical depression used self-report rather than performance based assessment tools, or assessed task performance on one or two levels of task performance rather than all three levels (overall, domain, and task). Furthermore,

none of the studies used a sample of only older women with major depression. Hence, for these reasons, we conducted a study (Chapter 4) to examine the influence of time on task performance in older women with major depression.

The second study (Chapter 3) examined the influence of the clinic and home on task performance in older women with major depression using a valid and reliable performance-based assessment tool, the Performance Assessment of Self-Care Skills (PASS). Using Rasch analysis (item difficulty logits and person ability logits) we compared task performance between the clinic and home on three levels of performance: overall, domain, and task. At the overall level, the results showed that task performance was significantly more independent at home than in the clinic. At the domain level, task performance was also significantly more independent at home that in the clinic for cognitive-oriented activities of daily living (CIADL) and physical-oriented activities of daily living (PIADL). Moreover, task performance was more independent at home than in the clinic for functional mobility (FM) and personal care activities of daily living (PC), but the results did not reach statistical significance. At the task levels, the results varied, with task performance being:

- significantly more independent at home for dressing, mail bills and checks, telephone use, obtain information: auditory, and clean up after meal preparation;
- more independent at home, but not statistically significant, for toilet transfer, indoor walking, oral hygiene, balance a checkbook, medication management, small repairs, home safety, play bingo, use sharp utensils, bending, lifting, carrying, and sweeping;
- significantly more independent in the clinic for stair use, bathtub/shower transfer, and oven use; and

- more independent in the clinic, but not statistically significant, for bed transfer, trimming toenails, obtaining info: visual, shopping, pay bills by checks, stovetop use, and change bed linens.

In brief, the number of tasks in which task performance was more independent at home was greater than the number that were more independent in the clinic. Hence, the results showed that the home was generally a more facilitative environment than the clinic for independent task performance for older women with major depression.

The third study, examined the effect of time on task performance in older women with major depression. Again we used Rasch analysis, item difficulty logits, and person ability logits to compare task performance over time on three levels of performance. At the overall level, our results showed that task performance tended to become less independent over time. At the domain level, task performance was significantly less independent for CIADL and PIADL. Finally, at the task level, the results also varied, with task performance being:

- significantly less independent over time for stovetop use and clean up after meal preparation;
- less independent over time, but not statistically significant, for stair use indoor walking, oral hygiene, bending, lifting, carrying, and sweeping;
- significantly more independent over time in home safety and bathtub/shower transfer;
- more independent over time, but not statistically significant, for toilet transfer and playing bingo; and
- stable for the rest of the tasks.

In brief, the largest number of tasks remained stable over time, however, the number in which task performance became less independent over time is greater than the number in which task performance became more independent over time. Hence, the results showed a general tendency of a less independent performance over time for older women with major depression.

In summary, the findings from this dissertation showed a significant influence of the environment and time on task performance in older women with major depression. The findings suggested that for older women with major depression task performance is better assessed in a familiar environment such as the home rather than an unfamiliar environment such as the clinic. Assessment in an unfamiliar environment may underestimate the patients' ability to perform daily activities. Moreover, older women with depression showed less overall independent performance over time as well as less independent performance in most functional domains and tasks assessed. Hence, attention should be given to those tasks for which performance was negatively affected by the environment or time. Consideration of these findings can affect intervention planning and implementation for this vulnerable population.

Future studies are needed to examine task performance in larger samples of older women with major depression, as well as older men with depression. Global measures of functioning and task specific measures of functioning should also be compared, given that our findings showed loss of independence in specific tasks (IADL) over only a 6 month time period. The influence of the environment, especially public environments, on task performance also needs further study to determine whether the impact of environment leads to overestimation or underestimation of performance in persons with depression. Finally,

future studies need to use more sophisticated and sensitive impairment measures, including measures of executive functioning, to better understand the relationship between impairment and task performance in persons with depression.

APPENDIX

THE RELIABILITY AND VALIDITY OF THE PASS

The interrater reliability of the PASS was established (Rogers et al., 2003) by administering the PASS to 57 older adults in a clinic (using the clinic version) and home (using the home version) and then calculating the percent of agreement (i.e., the number of agreements/ number of possible agreements) between two raters (from different staffs) in the four PASS domains. Interrater reliability was established for each version of the PASS and each domain. For the PASS-clinic, the percent of agreement between the raters was 99% for the 5 FM tasks, 98% for the 3 PC tasks, 94% for the 14 CIADL tasks, and 97% for the 4 PIADL tasks. For the PASS home, the percent of agreement between the raters was 97% for the 5 FM tasks, 91% for the 3 PC tasks, 93% for the 14 CIADL tasks, and 94% for the 4 PIADL tasks. These results indicate that the PASS has excellent interrater reliability. Test-retest reliability was also examined on two consecutive days with $r = .96$ which indicates excellent stability of the PASS.

Content validity of the PASS was established in reference to other functional assessments, the OARS Multidimensional Functional Assessment questionnaire-Activities of Daily Living (Fillenbaum, 1988; Pfeiffer, 1976), the rating scale for Physical Self-Maintenance and Instrumental Self-Maintenance (Lawton, & Brody, 1969; Lawton, Moss, Fulcomre, & Kleban, 1982), and the Functional Assessment Questionnaire (Pfeiffer, 1987;

Pfeiffer, Kurosaki, Harrah, Chance, & Filos, 1982) (see Appendix). Domains, tasks, and subtasks were reviewed in all these instruments. Then, performance tasks and subtasks were developed for each domain in the PASS. Construct validity of the PASS was established by administering the PASS to 6 samples: well older adults, older adults with osteoarthritis, cardiopulmonary diseases, depression, dementia, and macular degeneration. Well older adults scored highest on the PASS followed by older adults with depression, osteoarthritis, cardiopulmonary diseases, dementia, and macular degeneration. Older adults with dementia or macular degeneration scored significantly worse than all other samples, and older adults with osteoarthritis and cardiopulmonary diseases scored significantly lower than well older adults. The results indicated good construct validity because dependency was ordered as expected from the nature of the underlying impairment.

BIBLIOGRAPHY

Alexopoulos, G., Vrontou, C., Kakuma, T., Meyers, B., Yound, R. C., Klausner, E., et al. (1996). Disability in geriatric depression. *American Journal of Psychiatry, 153,* 877-885.

American Psychiatric Association. (2000). Diagnostic *and statistical manual of mental disorders* (4th ed., text revision). Washington, DC: Author.

Andrews, K., & Stewart, J. (1979). Stroke recovery: He can but does he? *Rheumatology and Rehabilitation, 18*(43), 43-48.

Arbuthnott, K., & Frank, J. (2000). Trail Making Test, part B as a measure of executive control validation using a set-switching paradigm. *Journal of Clinical and Experimental Neuropsychology, 22,* 518-528.

Armenian, H., Pratt, L., Gallo, J., & Eaton, W. (1998). Psychopathology as a predictor of disability: A population-based follow-up study in Baltimore, Maryland. *American Journal of Epidemiology, 148,* 269-275.

Bartels, S., & Smyer, M. (2002). Mental disorders of aging: An emerging public health crisis. *Generations: Journal of the American Society of Aging, 26*(1), 14-20.

Beekman, A., Deeg, D., Braam, A., Smit, J., & Van Tilburg, W. (1997). Consequences of major and minor depression in later life: A study of disability, well-being and services utilization. *Psychological Medicine 27,* 1397-1409.

Benjamin, J. (1976). The Northwick Park A.D.L. Index. *Journal of Occupational Therapy 39,* 301-306.

Berkman, L., Berkman, C., Kasl, S., Freeman, D., Leo, L., Ostfeld, A., et al. (1986). Depressive symptoms in relation to physical health and functioning in the elderly. *American Journal of Epidemiology, 124,* 372-388.

Blazer, D. (2003). Depression in late life: Review and commentary. *Journal of Gerontology, 58A,* 249-265.

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

Bottari, C., Dutil, E., Dassa, C., & Rainville, C. (2006). Choosing the most appropriate environment to evaluate independence in everyday activities: Home or clinic? *Australian Occupational Therapy Journal, 53,* 98-106.

Bowling, A., Farquhar, M., & Grundy, E. (1996). Outcome of anxiety and depression at two and a half years after baseline interview: Associations with changes in psychiatric morbidity among three samples of elderly people living at home. *International Journal of Geriatric Psychiatry, 11*(2), 119-129.

Brines, J. (1994). Economic dependency, gender, and the division of labor at home. *The American Journal of Sociology, 100,* 625-688.0

Bruce, M. (1999). The association between depression and disability. *American Journal of Geriatric Psychiatry, 7*(1), 8-11.

Bruce, M. (2000). Depression and disability In G. Williamson, D. R. Shaffer & P. Parmelee (Eds.), *Physical Illness and Depression in Older Adults: A Handbook of Theory, Research, and Practice* (pp. 11-29). New York, NY: Springer.

Bruce, M., Seeman, T., Merrill, S., & Blazer, G. (1994). The impact of depressive symptomatology on physical disability: MacArthur studies of successful aging. *American Journal of Public Health, 84*, 1796-1799.

Callahan, C. M., Hendrie, H. C., Dittus, R. S., Brater, D. C., Hui, S. L., & Tierney, W. M. (1994). Improving treatment of late life depression in primary care: A randomized clinical trial. *Journal of American Geriatrics Society, 42*, 839-846.

Callahan, C. M., Kroenke, K., Counsell, S., Hendrie, H. C., Perkins, A. J., Katon, W., et al. (2005). Treatment of depression improves physical functioning in older adults. *Journal of American Geriatrics Society, 53*, 367-373.

Callahan, C., Wolinsky, F., Stump, T., Nienaber, N., Hui, S., & Tierney, W. (1998). Mortality, symptoms, and functional impairment in late-life depression. *Journal of General International Medicine, 13*, 746-752.

Christiansen, C., & Baum, C. (1991). *Occupational Therapy: Overcoming human performance deficits*. Thorofare, NJ: Slack.

Chisholm, D. (2005). Disability in older adults with depression. Dissertation Abstract International 67 (01) (UMI No. 3206763)

Chou, K.-L. (2005). Everyday competence and depressive symptoms: Social support and sense of control as mediators or moderators? *Aging and Mental Health, 9*, 177-183.

Cristian, A. (2006). The assessment of older adult with a physical disability: A guide for clinicians. *Clinics in Geriatric Medicine, 22*, 221-238.

Czaja, S. J., Weber, R. A., & Nair, S. N. (1993). A human factor analysis of ADL activities: A capability-demand approach. *Journal of Gerontology, 48*, 44-48.

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.

Davison, M. (1991). Assessing environmental factors. In C. Christiansen & C. Baum (Eds.), *Occupational therapy: Overcoming human performance deficits*. Thorofare, NJ: Slack.

Desai, H., & Jann, M. (2000). Major depression in women: A review of the literature. *Journal of American Pharmaceutical Association, 40*, 525-537.

Druss, B., Rohrbaugh, R., & Rosenheck, R. (1999). Depressive symptoms and health costs in older medical patients. *American Journal of Psychiatry, 156*, 477-479.

Dunlop, D., Hughes, S., & Manheim, L. (1997). Disability in activities of daily living: Patterns of change and a hierarchy of disability. *American Journal of Public Health, 87*(3), 378-383.

Dunlop, D., Manheim, L., Song, J., Lyons, J., & Chang, W. (2005). Incidence of disability among preretirement Adults: The impact of depression. *American Journal of Public Health, 95*, 2003-2008.

Eberl, D. R., Fasching, V., Rahlfs, V., Schleyer, I., & Wolf, R. (1976). Repeatability and objectivity of various measurements in rheumatpid arthritis. *Arthritis and Rheumatism, 19*, 1278-1286.

Everson-Rose, S., Skarupski, K., Bienias, J., Wislson, r., Evans, D., & Mendes De Leon, C. (2005). Do depressive symptoms predict declines in physical performance in an elderly biracial population? *Psychosomatic Medicine, 67*, 609-615.

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

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

Fisher, A. (1995). The assessment of motor and process skills, Fort Collins, Colorado: Three Star Press.

Forsell, Y., Jorm, A., & Winblad, B. (1994). Association of age, sex, cognitive dysfunction and disability with major depressive symptoms in an elderly sample. *American Journal of Psychiatry, 151*, 1600-1604.

Forsell, Y., Jorm, A., & Winblad, B. (1998). The outcome of depression and dysthymia in a very elderly population: Results from a three-year follow-up study. *Aging and Mental Health, 2*, 100-104.

Fried, T., Bradley, E., Williams, C., & Tinetti, M. (2001). Functional disability and health care expenditures for older persons. *Archives of International Medicine, 161*, 2602-2607.

Fried, L., & Guralink, J. (1997). Disability in older adults: Evidence regarding significance, etiology, and risk. *Journal of American Geriatrics Society, 45*(1), 92-100.

Gallo, J., & Lebowitz, B. (1999). The epidemiology of common late-life mental disorders in the community: Themes for the new century. *Psychiatric Services, 50*, 1158-1166.

Gallo, J., Rabins, P., Lyketsos, C., Tien, A., & Anthony, J. (1997). Depression without sadness: Functional outcomes of nondysphoric depression in later life. *Journal of American Geriatric Society, 45*, 570-578.

Gallo, J., Rebok, G., Tennstedt, S., Wadley, V., Horgas, A., & The advanced cognitive training for independent and vital elderly (ACTIVE) study investigators (2003). Linking depressive symptoms and functional disability in late life. *Aging and Mental Health, 7*, 469-480.

Gärling, T., Böök, A., & Lindberg, E. (1985). Adults' memory representations of their everyday physical environment. In R. Cohen (Ed.), *The development of spatial cognition* (pp. 141-184). Hillsdale, NJ: Erlbaum.

Gill, T., Allore, H., Holford, T., & Guo, Z. (2004). Hospitalization, restricted activity, and the development of disability among older persons. *Journal of American Medical Association, 292*, 2115-2124.

Giovagnoli, A. R., Del Pesce, M., Mascheroni, M., Simoncelli, S., Laiacona, M., & Capitani, E. (1996). Trail Making Test: Normative values from 287 normal adult controls. *The Italian Journal of Neurological Sciences, 17*, 305-309.

Greelings, S., Beekman, A., Deeg, D., Twisk, J., & van Tilburg, W. (2001). The longitudinal effect of depression on functional limitation and disability in older adults: An eight-wave prospective community-based study. *Psychological Medicine, 31*, 1361-1371.

Guaralink, J., Leveille, S. G., Hirsch, R., Ferrucci, L., & Fried, L. (1997). The impact of disability in older women. *Journal of American Medical Women's Association, 52*(3), 113-119.

Guccione, A., Felson, D., Anderson, J., Anthony, J., Zhang, Y., Wilson, P., et al. (1994). The effects of specific medical conditions on the functional limitations of elder in the Framingham Study. *American Journal of Public Health, 84*, 351-358.

Gurland, B., Cross, P. S., & Katz, S. (1996). Epidemiological perspectives on opportunities for treatment of depression. . *American Journal of Geriatric Psychiatry, 4 (Supp. 1)*, S7-S13.

Gurland, B., Wilder, D., & Berkman, C. (1988). Depression and disability in the elderly: Reciprocal relations and changes with age. *International Journal of Geriatric Psychiatry, 3,* 163-179.

Han, B. (2002). Depressive symptoms and self-rated health in community-dwelling older adults: A longitudinal study. *Journal of American Geriatric Society, 50,* 1549-1556.

Han, B., & Jylha, M. (2006). Improvement in depressive symptoms and changes in self-rated health among community-dwelling disabled older adults. *Aging and Mental Health, 10,* 599-605.

Haworth, R., & Hollings, E. (1979). Are hospitals assessments of daily living activities valid? *International Rehabilitation Medicine, 1*(2), 59-62.

Hays, J., Saunders, W., Flint, E., Kaplan, B., & Blazer, D. (1997). Social support and depression as risk factors for loss of physical function in late life. *Aging and Mental Health, 1,* 209-220.

Hays, J., Steffens, D., Flint, E., Bosworth, H., & George, L. (2001). Does social support buffer functional decline in elderly patients with unipolar depression? *American Journal of Psychiatry, 158,* 1850-1855.

Hays, R., Wells, K., Sherbourne, C., Rogers, W., & Spritzer, K. (1995). Functioning and well-being outcomes of patients with depression compared with chronic general medical illnesses. *Archives of General Psychiatry, 52*(1), 11-19.

Heikkinen, R.-L., Berg, S., Avlund, K., & Törmakängas, T. (2002). Depressed mood: Changes during a five-year follow-up in 75-year-old men and women in three Nordic localities. *Aging: Clinical Experimental and Research, 14*(Suppl. 3), 16-28.

Heikkinen, R.-L., & Kauppinen, M. (2004). Depressive symptoms in late life: A 10-year follow-up. *Archives of Gerontology and Geriatrics, 38,* 239-250.

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

Holm, M. B., Rogers, J., & James, A. (2003). Interventions for daily living. In E. Blesedell Crepeau, E. Cohn & B. Boyt Schell (Eds.), *Willard & Spackman's Occupational Therapy* (10th ed, pp. 491-554). Philadelphia, Lippincott: Williams & Wilkins.

Katz, I. (1996). On the inseparability of mental and physical health in aged persons. *American Journal of Geriatric Psychiatry, 4,* 1-16.

Kennedy, G., Kelman, H., & Thomas, C. (1990). The emergence of depressive symptoms in late life: The importance of declining health and increasing disability. *Journal of Community Health, 15(2),* 93-104.

Kessler, R. C., Berglund, P., Demler, O., Jin, R., Koretz, D., Merikangaz, K., et al. (2003). The epidemiology of major depressive disorder *Journal of the American Medical Association, 289*, 3095-3105.

Kessler, R. C., Chiu, W. T., Demler, O., & Walters, E. E. (2005). Prevalence, severity, and comorbidity survey replication (NCS-R). *Archives of General Psychiatry, 62*, 617-627.

Kiosses, D., & Alexopoulos, G. (2005). IADL functions, cognitive deficits, and severity of depression: A preliminary study. *American Journal of Geriatric Psychiatry, 13*, 244-249.

Kurland, B., Gill, T., Patrick, D., Larson, E., & Phelan, E. (2006). Longitudinal change in positive affect in community-dwelling older persons. *Journal of American Geriatric Society, 54*, 1846-1853.

Lachs, M., Feinstein, A., Cooner, L., Drickamer, M., Marottoli, R., Pannill, F., et al. (1990). A simple procedure for screening functional disability in elderly patients. *Annals of Internal Medicine, 112*, 699-706.

Laukkanen, P., Era, P., Heikkinen, R.-L., Suutama, T., Kauppinen, M., & Heikkinen, E. (1994). Factors related to carrying out everyday activities among elderly people aged 80. *Aging: Clinical Experimental and Research, 6*, 433-443.

Laukkanen, P., Heikkinen, E., Schroll, M., & Kauppinen, M. (1997). A comparative study of factors related to carrying out physical activities of daily living (PADL) among 75-year-old men and women in two Nordic localities. *Aging: Clinical Experimental and Research, 9*, 258-267.

Lawton, M. P. (1982). Competence, environmental press, and the adaptation of older people. In M.P. Lawton, P.G. Windley, & T.O. Byerts (Eds), *Aging and the environment* (pp. 33-59). New York, NY: Springer Publishing Company.

Lawton, M. P. (1983). Environment and other determinants of well-being in older people. *The Gerontologist, 23*, 349-356.

Lawton, M. P. (1986). *Environment and aging* (2nd ed.). Albany, NY: Plenum.

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

Lawton, M. P., Moss, M., Fulcomer, M., Kleban, M.H. (1982). A research and services oriented multilevel assessment instrument. *Journal of Gerontology*, 37, 91-99.

Lenze, E. J., Rogers, J. C., Martire, L. M., Mulsant, B. H., Rollman, B. L., Dew, M. A., et al. (2001). The association of late-life depression and anxiety with physical disability. *American Journal of Geriatric Psychiatry 9*, 113-135.

Lenze, E. J., Schulz, R., Marirel, L. M., Zdaniuk, B., Glass, T., Kop, W., et al. (2005). The course of functional decline in older people with persistently elevated depressive symptoms: Longitudinal findings from the Cardiovascular Health study. *Journal of American Geriatrics Society, 53*, 569-575.

Linacre, M. (2006). WINSTEPS Rasch measurement computer program (Version 3.63.2) [Computer software]. Chicago, IL.

Lopezs, A., & Murray, C. J. (1998). The global burden of disease, 1990-2020. *Nature Medicine, 4*, 1241-1243.

Marini, M., & Shelton, B. (1993). Measuring household work: Recent experience in the United States. *Social Science Research, 22*, 361-382.

McDowell, I., Kristjansson, B., Hill, G. B., & Hebért, R. (1997). Community screening for dementia: The Mini Mental State Exam (MMSE) and Modified Mini-Mental State Exam (3MS) compared. *Journal of Clinical Epidemiology, 50*, 377-383.

Mirowsky, J., & Ross, C. (1992). Age and depression. *Journal of Health and Social Behavior, 33*(3), 187-205.

Mojtabai, R., & Olfson, M. (2004). Cognitive deficits and the course of major depression in a cohort of middle-aged and older community-dwelling adults. *Journal of American Geriatrics Society, 52*, 1060-1069.

Mulsant, B., Ganguli, M., & Seaberg, E. (1997). The relationship between self-rated health and depressive symptoms in an epidemiological sample of community-dwelling older adults. *Journal of American Geriatric Society, 45*, 954-958.

Murray, C. J., & Lopez, A. D. (1996). *The global burden of disease. A comprehensive assessment of mortality and disability from diseases. Injuries and risk factors in 1990 and projected to 2020*. Cambridge, MA: Harvard University press.

National Institute of Health. (1992). "NIH consensus development conference statement: Diagnosis and treatment of depression in late life. *Journal of the American Medical Association, 268*, 1018-1024

National Institute of Mental Health. (2007, September). Retrieved from http://www.nimh.nih.gov/health/publications/older-adults-depression-and-suicide-facts.shtml

National Institute of Mental Health (2007, November). Major Depression. Retrieved from http://www.nimh.nih.gov/health/topics/depression/index.shtml

Nichols, P. (1976). Are ADL indices of any value? *British Journal of Occupational Therapy, 39*, 160-163.

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

Ormel, J., Kempen, G. I., Deeg, D., Brilman, E., van Sonderen, E., & Reyveld, J. (1998). Functioning, well-being, and health perception in late middle-aged and older people: Comparing the effects of depressive symptoms and chronic medical conditions. *Journal of American Geriatric Society, 46*, 39-48.

Ostir, G., Volpato, S., Kasper, J., Frerrucci, L., & Guralnik, J. (2001). Summarizing amount of difficulty in ADL: A refined characterization of disability. Results from the Women's Health and Aging Study. *Aging: Clinical Experimental and Research, 13*, 465-472.

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. *The American Journal of Occupational Therapy, 48*, 697-709.

Penninx, B., Leveille, S., Ferrucci, L., van Eijk, J., & Guralnik, J. (1998). Exploring the effect of depression on physical disability: Longitudinal evidence from the established populations for epidemiologic studies of the elderly. *American Journal of Public Health, 89*, 1346-1352.

Pfeiffer, E. (1976) Multidimensional functional assessment: The OARS methodology. Durham, NY: Center for the Study of Aging and Development.

Pfeffer, R.I. (1987). The Activities Questionnaire. In I. McDowell & C. Newell (Eds.), Measuring Health: A guide to rating scales and questionnaires. New York, NY: Oxford Press.

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

Prince, M., Harwood, R., Blizard, R., Thomas, A., & Mann, A. (1997). Impairment, disability and handicap as risk factors for depression in old age. The Gospel Oak Project V. *Psychological Medicine, 27*, 311-321.

Prince, M., Harwood, R., Thomas, A., & Mann, A. (1998). A prospective population-based cohort study of the effects of disablement and social milieu on the onset and maintenance of late-life depression. The Gospel Oak Project VII. *Psychological Medicine, 28*, 337-350.

Portney, L., & Watkins, M. (2000). *Foundations of clinical research: Applications to practice.* (2nd ed). Upper Saddle River, NJ: Prentice Hall Health.

Raina, K., Rogers, J., & Holm, M. (2007). Influence of the environment on activity performance in older women with heart failure. *Disability and Rehabilitation, 29*, 545-557.

Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Copenhagen, Denmark: Danmarks Paedagogiske Institute.

Rigby, P., & Letts, L. (2003). Environment and occupational performance: Theoretical considerations. In L. Lori, P. Rigby & D. Stewart (Eds.), *Using environments to enable occupational performance* (pp. 17-32). Therofare, NJ: SLACK Incorporated.

Rogers, J. C., & Holm, M. B. (1989). Performance Assessment of Self-Care Skills. Unpublished performance test. University of Pittsburgh; 1989.

Rogers, J. C., & Holm, M. B. (2000). Daily-living skills and habits of older women with depression. *Occupational Therapy Journal of Research, 20 (Suppl.),* 68s-85s.

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

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

Rogers, J., Holm, M., Goldstein, G., McCue, M., & Nussbaum, P. (1994). Stability and change in functional assessment of patients with geropsychiatric disorders. *American Journal of Occupational Therapy, 48*, 914-918.

Schroll, M., Avlund, K., & Davidsen, M. (1997). Predictors of five-year functional ability in a longitudinal survey of men and women aged 75 to 80. The 1914-population in Glostrup, Denmark. *Aging: Clinical Experimental and Research, 9*, 143-152.

Schulberg, H., Madonia, M., Block, M., Coulehan, J., Scott, C., Rodriguez, E., et al. (1995). Major depression in primary care practice: Clinical characteristic and treatment implications. *Psychosomatics, 36*, 129-137.

Seaburn, D., Lyness, J., Eberly, S., & King, D. (2005). Depression, perceived family criticism, and functional status among older, primary-care patients. *American Journal of Geriatric Psychiatry, 13*, 766-772.

Sheehan, D., V. (2004). Depression: Underdiagnosed, undertreated, underappreciated. *Managed Care, 13*, 6, 6-8.

Sheikh, K., Smith, D., Goldenberg, E., Brennan, P., & Kinsella, G. (1979). Repeatability and validity of a modified activities of daily living (ADL) index in studies of chronic disability. *International Rehabilitation Medicine 1*(2), 51-58.

Shelton, B., & John, D. (1996). The division of household labor. *Annual Reviews of Sociology, 22*, 299-322.

Shumway-Cook, A., Patla, A. E., Stewart, A., Ferrucci, L., Ciol, M. A., & Guralnik, J. M. (2002). Environmental demands associated with community mobility in older adults with and without mobility disabilities. *Physical Therapy, 82,* 670-681.

Simon, G., Katon, W., Rutter, C., VonKroff, M., Lin, E., Robinson, P., et al. (1998). Impact of improved depression treatment in primary care on daily functioning and disability. *Psychological Medicine, 28,* 693-701.

Spector, W. (1996). Functional disability scales. In B. Spilker (Ed.), *Quality of Life and Pharmacoeconomics in Clinical Trials* (pp. 133-143). Philadelphia, PA: Lippincott-Raven.

SPSS Inc. (2002). (Version 15.0). Chicago, IL.

Steffens, D., Hays, J., & Krishnan, K. (1999). Disability in geriatric depression. *American Journal of Geriatric Psychiatry, 7,* 34-40.

Stuck, A., Walthert, J., Nikolaus, T., Büla, C., Hohmann, C., & Beck, J. (1999). Risk factors for functional status decline in community-living elderly people: A systematic literature review. *Social Science and Medicine, 48,* 445-469.

Teng, E., & Chui, H. (1987). The Modified Mini-Mental State (3MS) examination. *Journal of Clinical Psychiatry, 48,* 314-318.

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

Tombaugh, T. (2004). Trail Making Test A and B: Normative data stratified by age and education. *Archives of Clinical Neuropsychology, 19,* 203-2014

Turner, R. J., & Noh, S. (1988). Physical disability and depression: A longitudinal analysis. *Journal of Health and Social Behavior, 29,* 23-37.

Unützer, J., Patrick, D., Simon, G., Grembowski, D., Walker, E., Rutter, C., et al. (1997). Depressive symptoms and the cost of health services in HMO patients aged 65 years and older. *Journal of the American Medical Association, 277,* 1618-1623.

U.S Department of Health and Human Services. (2006). A statistical profile of older Americans aged 65+. Retrieved from www.aoa.gov/press/fact/pdf/ss_stat_profile.pdf

U.S Department of Health and Human Services. (2006). A statistical profile of older Americans aged 65+. Retrieved from www.aoa.gov/prof/Statistics/profile/ 2006/profiles2006.asp

Vaccarino, V., Kasl, S., Abramson, J., & Krumholz, H. (2001). Depressive symptoms and risk of functional decline and health in patients with heart failure. *Journal of American College of Cardiology, 38,* 199-205.

Wells, K., Burnam, M. A., Rogers, W., Hays, R., & Camp, P. (1992). The course of depression in adult outpatients. *Archives of General Psychiatry, 49*, 788-794.

Wells, K. B., Stewart, A., Hays, R. D., Burnam, A., Rogers, W., Daniels, M., et al. (1989). The functioning and well-being of depressed patients. *Journal of American Medical Association 262*, 914-919.

West, S. K., Rubin, G. S., Munoz, B., Abraham, D., Fried, L., & Salisbury Eye Evaluation Project Team. (1997). Assessing functional status: Correlation between performance on tasks conducted in a clinic setting and performance on the same task conducted at home. *Journal of Gerontology, 52A*, M209-M217.

Williamson, G., & Schulz, R. (1992). Physical illness and symptoms of depression among elderly outpatients. *Psychology and Aging, 7*, 343-351.

Whyte, E. M., Mulsant, B. H., Vanderbilt, J., Dodge, H. H., & Ganguli, M. (2004). Depression after stroke: A prospective epidemiological study. *Journal of American Geriatrics Society, 42*, 774-778.

Yerxa, E., & Baum, S. (1987). Environmental theories and the older person. *Topics in Geriatric Rehabilitation, 3*(1), 7-18.

Yesavage, J., Brink, T., Rose, T., Lum, O., Huang, V., Adey, M., et al. (1983). Development and validation of a geriatric depression screening scale: A preliminary report. *Journal of Psychiatric Research, 17*, 37-49.

Zeiss, A., Lewinsohn, P., Rohde, P., & Seeley, J. (1996). Relationship of physical disease and functional impairment to depression in older people. *Psychology and Aging, 11*, 572-581.