Evaluation of functional fitness status and examination of its underlying structure in older adults

by

Katerina Machacova

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Kinesiology)

Charles University
Faculty of Physical Education and Sports
Sport Research Center

2009

Mentor: Prof. Ing. Vaclav Bunc, CSc.

Advisors: Diane Adamo, Ph.D. and MUDr. Iva Holmerova, Ph.D.

Written at:

Wayne State University
Institute of Gerontology
Detroit, Michigan - USA

Abstract

Quality of life in later years depends to a large degree on the ability to function independently which in other words means to have a capacity to perform daily tasks without pain, undue fatigue and for as long as needed. The functional fitness status is the most important presumption of the independency. It is essential from the point of view of society to help older adults to maintain certain fitness level thus maintain their independency for as long as possible. Therefore, an accurate evaluation of relevant fitness components is crucial but challenging at the same time. Many factors make this population difficult to test. Instruments should be sensitive enough to detect even small but clinically significant changes, safe, and quick-and-easy to perform by the majority of older individuals, preferably in the field setting.

In the Czech Republic there is a lack of relevant information about older population's fitness measured by objective tests which are increasingly used by researchers and clinicians all around the world. In past few decades many instruments that were standardized to measure physical functioning among older adults have been developed. Hence, the first general aim of this dissertation was to overview and to organize existing assessment methods of fitness evaluation and to discuss the most commonly used instruments. The second aim was to apply the most sufficient instrument on older Czech adults in order to gain an experience with performance testing and to provide baseline data which would help to shape future research. The last general aim was more theoretical. The problem of composite measurement score development when measuring latent constructs was addressed. Psychometric evaluation of scales which measure unobservable constructs continues to be an issue of high interest among many researchers. However, current practice showed serious flaws resulting from a lack of basic measurement properties and requirements. This last study will not only answer questions what is the most accurate approximation of functional fitness and what is the contribution of each component to the overall score but will also provide guidelines for any composite measurement score development in relevant areas of behavioral research and beyond.

The present dissertation was divided into three independent but interrelated studies each of which was addressed to earlier defined general aims. Overall findings, conclusions, and some practical implications are then conferred in the general discussion.

Abstrakt v českém jazyce

Kvalita života v pozdním věku do značné míry závisí na schopnosti nezávisle vykonávat aktivity každodenního života a to sice bez bolesti, nadbytečné únavy a po potřebně dlouhou dobu. Funkční zdatnost je jedním z nejzákladnějších předpokladů zmíněné nezávislosti. Z pohledu celé společnosti je velmi důležité pomáhat lidem v pokročilém věku udržet si minimální nutnou úroveň funkční zdatnosti nutné právě k samostatnému vykonávání každodennínch aktivit a to po co nejdelší dobu. S tím úzce souvisi potřeba schopnosti co nejpřesněji měřit jednotlivé komponenty zmíněné zdatnosti. Měření osob v pokročilém věku ale skýtá určité problémy. Mnoho faktorů totiž činí testovnáni starších osob velmi komplikovaným úkolem. Vhodný instrument by měl být dostačně sensitivní k odhalení i malých, leč klinicky významných změn, ale přesto bezpečný a snadno proveditelný pro valnou většinu populace nad 60 let. Navíc je vhodné, aby bylo dané testování proveditelné v terénních podmínkách.

V České republice je evidentní nedostatek kvalitních informací o fyzické způsobilosti seniorské populace. Dá se říci, že zkušenost s objektivními testy založenými na kvantitativním hodnoceni provádění daných úkonů nebyla prakticky systematicky zdokumentována. V zahraničí v poslední době rapidně stoupá obliba používání zmíněných objektivních metod hodnocení zdatnosti. S tim souvisí i narůstajicí počet nejrůznějších standardizovaných testů či testových baterii. Z tohoto důvodu prvnim cilem této disertace bylo identifikovat nejčastěji používané testy a testové baterie a přehledně je uspořádat dle způsobu hodnocení funkční zdatnosti. Celkem logicky následoval druhý obecný cíl, který se týkal aplikace nejvhodnější testové baterie na českou seniorskou populaci. Získané zkušenosti spolu s naměřenými daty mohou významně přispět ke směřování budoucího výzkumu v této relativně nové oblasti. Poslední obecný cíl byl zaměřen vice teoreticky. Konkrétně se týkal problému s měřením latentnich konstruktů, jakým je právě například funkční zdatnost. Psychometrické měření těchto konstruktů je stále častým zájmem mnoha výzkumných týmů. Přesto po podrobné rešerži odborné literatury byly identifikovány zásadní nedostatky v konstrukci zmíněných kompozitních skórů. Tato posledni část disertace nezodpoví pouze otázku, jak nejpřesněji měřit, respektive odhadovat celkovou funkční zdatnost, ale i jaká je relativní důležitost jednotlivých komponent vzhledem k

celkovému skóru. Zároveň může sloužit jako model pro měření jiných latentnich konstruktů, které jsou obzvláště v oblasti behaviorálního výzkumu tak běžné.

Tato disertace je rozdělena do třech na první pohled samostatných, ale jinak úzce souvisejících studií. V každé studii je detailněji rozpracován jeden z výše zmíněných obecných cílů. Obecné závěry jsou pak okomentovány a shrnuty v závěrečné diskusi.

Acknowledgement

I would like to acknowledge my mentor Vaclav Bunc for his distant guidance and feedback throughout whole dissertation process.

I would like to thank Mark Luborsky for providing me with a great opportunity to be part of the Institute of Gerontology. This would not be possible without Peter Lichtenberg, who is the director of this institute. I would also like to thank Cathy Lysack for her great ideas and motivation. And importantly, I am most grateful for voluntary mentorship of Diane Adamo. It has been great help through the process.

I would also like to acknowledge Iva Holmerova and Hana Vankova, who have invited me to participate on large study sponsored by Ministry of Health of the Czech Republic focused on effects of dance on different aspects of quality of life in older adults.

I am also deeply indebted to Pert Blahus who has given me the inspiration and encouragement over years.

Finally, I would like to acknowledge foundation "Nadani Jesefa, Marie a Zdenky Hlavkovych" and "Fond Mobility UK" for providing me with a financial support.

This is the confirmation that the present dissertation has been written independently by the author and that all used resources were properly cited.

TABLE OF CONTENTS

List of tables

List of figures

List of symbols

AADL	Advanced Activities of Daily Living
AAHPERD	American Alliance for Health, Physical Education, Recreation, and Dance battery
ADL	Activity of Daily Living
BADL	Basic Activities of Daily Living
BMA	Basic Motor Ability
CFA	Confirmatory factor analysis
CMS	Composite measurement score
CS-PFP	Continuous Scale Physical Functional Performance battery
EPESE	Established Populations for Epidemiologic Studies of the Elderly
FAM	Functional Assessment Measure
FFB	Functional Fitness Battery
FIM	Functional Independence Measure
GFE	Groningen Fitness test for the Elderly
GFI	Goodness of fit indices
IADL	Instrumental Activities of Daily Living
NFI	Normed fit index
PDI	Physical Disability Index
PFP-10	Physical Functional Performance 10

PPME	Physical Performance and Mobility Examination battery
PPT	Physical Performance Test battery
RMR	Root mean square residuals
RMSEA	Root mean square error of approximation
ROM	Range of motion
SEM	Structural Equation Modeling
SF-36	Short Form 36 questionnaire
SFT	Senior Fitness Test
SPPB	Short Physical Performance Battery
SPSS	Statistical Package for the Social Sciences

CHAPTER 1 – Introduction

1.1 Aging, physical activity and functional fitness

Aging is an experience that every human being shares but not fully understands. Although all people age, they do it in different ways and at different rates. One of the factors that make the study of aging so difficult is that a typical older adult does not exist. The term 'aging' refers to a process or a group of processes occurring in living organisms that with the passage of time lead to a loss of the adaptability, functional impairment, and eventually death. Aging is a logical extension of the physiological processes of growth and development, beginning with birth. Aging occurs with the relentless march of time, but relatively few people actually die of old age. Most die because the body loses the capacity to withstand physical or environmental stressors.

The goal of gerontologists and applied health scientists is to change the shape of the human survival curve so that most individuals can live longer lives. Several controllable factors, such as food restriction and nutrition (Kirk, 2001; Mobbs et al., 2001; Poehlman et al., 2001), general activity level (DiPietro, 2001), and physical activity (Blyth, Cumming, Mitchell, & Wang, 2007; Daley & Spinks, 2000; Ferrucci & Simonsick, 2006; Fiantore Singh, 2002; Frankel, Bean, & Frontera, 2006; Hollmann, Struder, Tagarakis, & King, 2007; Shepard, 1997; Stewart, 2005) have some promise in fulfilling that goal. However, most people would agree that a long life without health and physical independency is undesirable, yet many live their terminal years in a state of morbidity, or complete physical dependence and poor health. Discussions of extending the life span should always be entangled with issues of quality of life.

Active life expectancy is a term coined by Katz and his colleagues (Katz et al., 1983) that combines mortality and disability data. Active life expectancy refers to the number of remaining years of life that an individual may expect to be able to conduct the basic activities of daily living (ADL). Individuals who cannot carry out ADL activities are in a state of morbidity, are dependent on others, and have a low quality of life by most people's standards. Quality of life in later years depends to a large degree on being able to do the things we want to do, without pain, and for as long as possible. The quality of life is highly modifiable by physical activity through affecting functional fitness status (de Vreede et al., 2007; Kruger, Bowles, Jones, Ainsworth, & Kohl, 2007; Sawatzky, Liu-Ambrose, Miller, & Marra, 2007; Shibata, Oka, Nakamura, & Muraoka, 2007; Tessier et al., 2007). In other words, to be able to do things one wants to do require a certain functional fitness level.

As we are living longer it is increasingly important to pay attention to our activity levels. Numerous of technological advances in recent years have had mixed benefits for people relative to quality and quantity of life. Whereas medical technology has contributed to a longer life expectancy, computer technology is resulting in increasing sedentary lifestyles and increasing risk for chronic health and morbidity problems. Many older adults due to their sedentary lifestyle are functioning dangerously close to their maximum ability level during normal activities of daily living. Climbing stairs or getting out of the chair, for example, often require almost maximum efforts for older people who are not physically active. Any further decline could easily cause them to move from independent to dependant status in which assistance is needed. An adequate fitness level is needed to maintain independent and self-supported life for as long as possible as illustrated in figure 1.1.

The decline in fitness is caused either by pathology or by age related changes in cardiovascular and pulmonary system, muscle strength and power, flexibility, postural control system, cognitive function, mental well-being etc. However, there is evidence to support the findings that participation in physical activity and exercise programs may mitigate declines in most of the previously mentioned systems and thus reduce the onset of several disease processes (Fox, 1999; Frankel et al., 2006; Hollmann et al., 2007; Kramer & Erickson, 2007) and maintain independency. On the other hand, lack of physical activity has been associated with increasing insulin resistance and decreasing lipoprotein lipase activity in the skeletal musculature that can lead to chronic diseases such as atherosclerosis, with follow-up effects like coronary heart disease, myocardial insufficiency, hypertension, stroke, and type II diabetes (Boyle, Buchman, Wilson, Bienias, & Bennett, 2007; DiPietro, 2001; Haskell et al., 2007; Sawatzky et al., 2007; Stewart et al., 2005). In addition, there are related conditions that can result from lack of activity such as obesity, several types of cancer, osteoporosis and sarcopenia. Furthermore, physical inactivity speeds up the aging process in many people whereas increased physical activity slows it down in others.

The necessity, from the point of view of every individual as well as of society as whole, is that all that factors that serve to maintain health and fitness and intellectual capacity in old age should be exploited. According to Fiantore Singh (2002) the rationale supporting older individuals participation in physical activity and exercise can be divided into four broad themes: minimizing physiological changes; increasing longevity and decreasing the risk of many chronic diseases; treating certain chronic diseases; and preventing and treating disability (Fiantore Singh, 2002). The beneficial effects of physical activity and exercise on various physiological parameters such aerobic endurance, muscle strength, flexibility, and balance in the elderly have been well established (Baker, Atlantis, & Fiatarone Singh, 2007; Brown et al., 2000; Buchman, Boyle, Wilson, Bienias, & Bennett, 2007; Cao, Maeda, Shima, Kurata, & Nishizono, 2007; Capodaglio, Capodaglio Edda, Facioli, & Saibene, 2007; Conn, Minor, Burks, Rantz, & Pomeroy, 2003; DiBrezzo, Shadden, Raybon, & Powers, 2005; Frankel et al., 2006; Haskell et al., 2007; Hauer, Becker, Lindemann, & Beyer, 2006; Judge, Lindsey, Underwood, & Winsemius, 1993; Paterson, Jones, & Rice, 2007; Stewart, 2005) and these physiological parameters are the foci of our attention in this study.

Analysis of the literature focusing on key exercise variables (e.g., intensity, type, and volume) suggests that the requisite beneficial amount of activity is that which produces improved aerobic endurance, muscle strength, flexibility, and indirectly balance. All of these are fundamental and unavoidable components of functional fitness. Age-related declines in these components are such that physical limitations impinge on functional fitness and of course on the performance of activities of daily living. However, an exercise program can minimize declines and prevent older adults from crossing functional thresholds of inability (Paterson et al., 2007). In other words it is important for older adults to maintain an adequate functional fitness level. The easiest way is to pay attention to physical activity even or especially in old age, and to participate in exercise programs designed to improve physical fitness components. To ensure the possible highest effectiveness of these programs a fitness status evaluation has to be done.

1.2 Role and importance of adequate functional fitness level

Many people with and without disabilities often fail to realize the importance of being physically fit for the unexpected events in life. Most of the time we do not have to extend ourselves beyond the minimum levels of strength and balance necessary for performing routine activities of daily living. But there are certain times when greater levels of strength and cardiorespiratory fitness are needed to overcome uncharacteristic events in the indoor or outdoor environment. The examples might be: maneuvering over an uneven sidewalk or up a steep ramp; increasing pace in order to reach an approaching bus; getting back up after a fall; maintaining balance against high winds; getting out of a car; standing in line an extra 5 minutes; dealing with high temperatures and humidity; or carrying a heavier bag.

Functional fitness is defined as having a physical capacity to perform normal everyday activities safely and independently without undue fatigue (Rikli & Jones, 2001). But having the capacity to perform independently does not ensure independency. There are other factors that play an important role such as health status (number of chronic conditions), cognitive functioning, sensomotoric functioning, motor control, or even environment. The combination of earlier mentioned factors will influence an ability to perform independently on daily basis as illustrated in figure 1.1. However, this study is limited to the initial and the most preventable factor – functional fitness with the special attention targeted on measurement issues.

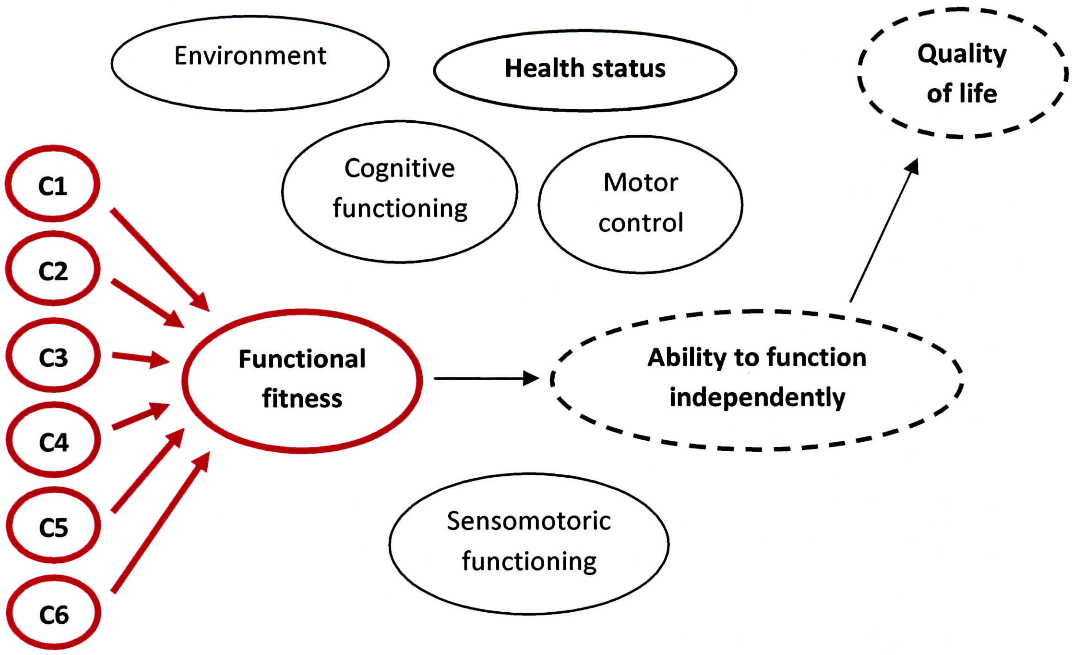


Figure 1.1 Diagram illustrating the focus of this study – how functional fitness contributes to the quality of life (C1 – C6 represents individual functional fitness components)

Physical activity plays an important role in maintaining functional fitness and is considered to be a part of the disability process (Nagi, 1965, 1991). The traditional models explaining the disabling process describe four main stages in the progression to disability: disease/pathology, physiological impairment, functional limitation, and disability. More specifically the model suggests that the pathology leads to physiological impairment (decline in body systems), physiological impairments leads to functional limitation (restriction in physical behavior), and functional limitation leads to disability (the inability to perform activities of daily living).

A

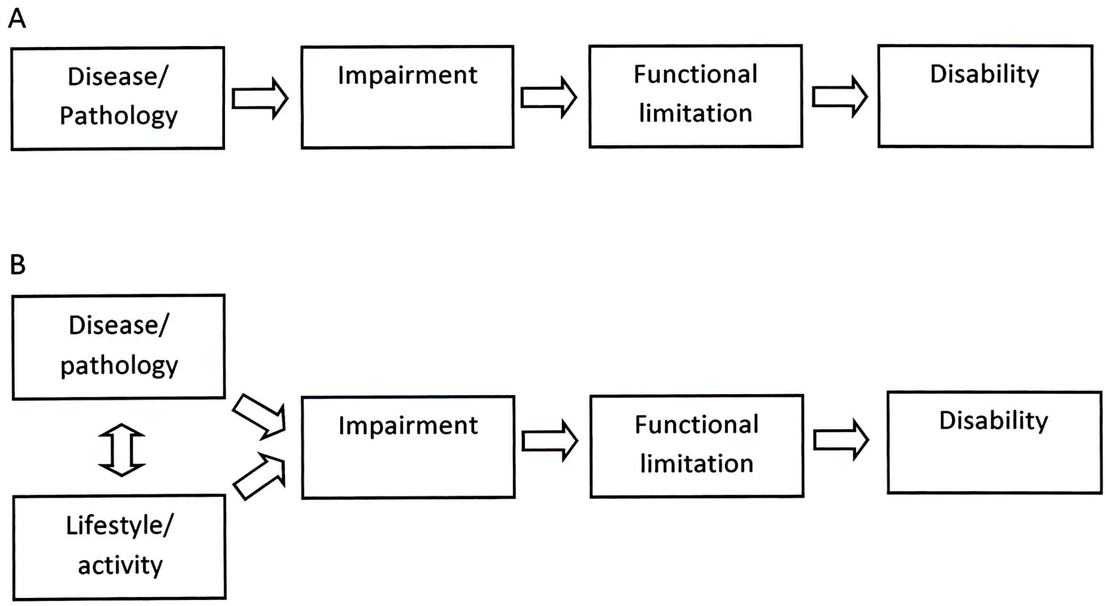


Figure 1.2 Disability model by Nagi et al. 1965 (A) and amended disability model by Nagi et al. 1991 (B)

Impairments refer to abnormalities at the level of tissues, organs, and body systems, whereas functional limitations are deficits in the ability to perform discrete tasks such as climbing stairs. Disability on the other hand refers to a functional limitation expressed in social context such as an inability to clean one's home or to shop independently. Although, it was traditionally thought that all disability originated from disease or pathology (figure 1.2 A), recent evidence suggests that a physically inactive lifestyle also can be a primary cause of frailty in later years especially for people living into their 80's and 90's, and that the traditional model should be amended as shown in figure 1.2 B. In fact, data suggests that physical inactivity is on a par with chronic disease as a cause of disability, and that increased physical activity is associated with higher levels of functional mobility in people who have chronic health problems as well as in those who are relatively healthy.

Very important is the evidence that physical decline, whether due to disuse or to disease, is modifiable through proper assessment and activity intervention. Understanding both contributing causes of physical decline and subsequent stages leading to frailty is helpful in planning effective prevention strategies and in designing sensitive measurement instruments.

1.3 Functional fitness components

The framework presented in table 1.1 illustrates the progressive relationship among functional fitness components, functional abilities, and activity goals. The common activities in the far right column (taking care of personal needs, household chores, shopping, or traveling) require the ability to perform the functions listed in middle column (walking, stair climbing, lifting, and reaching). These functions, in turn, require the adequate reserve in fitness components identified in the left column (muscular strength, aerobic endurance, flexibility, and agility/dynamic balance).

Table 1.1 Framework of functional fitness by Rikli and Jones (2001)

Functional fitness components	**Actual function**	**Activity goals**
Muscle strength	Walking	Personal care
- Upper-body	Stair climbing	Shopping or errands
- Lower-body	Standing up from the chair or bed	Housework
Aerobic endurance	Lifting or reaching	Gardening
Flexibility	Bending or kneeing	Travelling
- Upper-body		Sports
- Lower-body		
Dynamic balance		

Based on the functional fitness framework and the empirical evidence, the following components were identified as being the relevant components of functional fitness: muscle strength, aerobic endurance, flexibility, agility/dynamic balance.

1.3.1 Muscular strength

Strength is defined as the instantaneous maximal force generated by a muscle or group of synergistic muscles at a given velocity of the movement. Strength is physiologically dependent on both the number and diameter of myofibrils within muscle cells, on the fiber type, and on the coordination of the neurologic elements that control the contraction of skeletal muscle (Frankel et al., 2006).

Strength is essential to individuals across the life span because it is needed for many activities of daily living. Much of the strength loss with age is attributed to a loss of muscle mass caused by equal decline in number of type I and type II fibers. Also, there is a decline in muscle fiber size, primarily in type II fibers. In addition, there are changes in the nervous system with age including a loss of motor units, particularly fast-twitch units, and a larger inervation ratio in the motor units that remain. Declines in blood flow may also explain some of the strength loss (Spirduso, 2005).

According to fitness experts, maintaining muscular strength should be a primary concern. The decline in muscle strength, which averages about 15 – 20% per decade after the age of 50, can have devastating effect on people's ability to perform normal everyday activities of daily living. Maintaining strength and muscle function is also important because of the role it plays in helping to reduce the risk for falls and fall-related injuries, and because of its positive effect on a number of health conditions. Muscular strength can help reduce bone loss, improve glucose utilization, maintain lean body tissue, and prevent obesity (Rikli & Jones, 2001).

1.3.2 Aerobic endurance

Endurance is the ability to maintain a given level of exercise over time or to perform a given task repeatedly without undue fatigue that further prevents such activity. This factor is rooted in numerous physiologic parameters: air exchange in the lungs, heart function, blood circulation and patency of blood vessels, and the biochemical characteristic of individual muscle cells. Disease or conditions such as coronary or peripheral vascular disease, restrictive or obstructive pulmonary disease, general deconditioning, and malnutrition may therefore negatively affect endurance (Frankel et al., 2006).

Aerobic endurance is covered by cardiovascular and respiratory systems. The capacity to perform physical work is dependent to a large extend on the proper functioning of these two systems. Some structural changes in the cardiovascular system with age include increased thickening and hardening of the walls of the blood vessels and left ventricular wall and increased stiffness in the arteries. Systolic and diastolic pressures increase with age, attributable primarily to a thickening and hardening of the aorta and arterial tree but also to an increase in total peripheral resistance. The heart and vasculature also become less sensitive to β-mediated stimulation, and thus the aging heart cannot achieve maximum heart rate levels that were possible in youth. The heart rates of older people also remain higher and recover more slowly after maximal exercise. Postural hypotension, which predisposes an individual to dizziness, confusion, weakness, or fainting, increases with age but appears to be more related to high levels of systolic blood pressure than to aging per se. More structural and functional deteriorations are attributed to pathological process, such as heart disease, rather than the actual aging (Spirduso, 2005).

Clearly, aerobic endurance is an important fitness component. An adequate level of aerobic endurance is necessary to perform many activities of daily living. Although aerobic capacity tends to decline at the rate of 5 – 15% per decade after the age of 30, resulting in as much as a 50% loss by the age of 70. Studies indicate that at least half of the decline could be avoided by being physically active. Maintaining an adequate level of aerobic activity has both a direct effect on person's functional mobility and an indirect effect through its role in helping to reduce the risk for such medical conditions as cardiovascular disease, diabetes, obesity, high blood pressure, and some forms of cancer (Rikli & Jones, 2001).

1.3.3 Flexibility

Flexibility is described as range of motion (ROM) around a joint or joints in the body. The extensibility and intactness of many structures contribute to flexibility, including joint articular surfaces and capsules, and the loss of a connective tissue around muscles and tendons themselves (Frankel et al., 2006).

The aging process results in decreased collagen synthesis in skin, ligaments, tendons, and underlying tissues, which may lead to slower healing and adaptation to changing movement patterns. Blood flow through these tissues and through muscle also decreases, further inhibiting the process. Prolonged bed rest or common neurologic impairments in the elderly population that weaken muscles or cause spasticity lead to contracture mechanisms. First, joint capsules and surrounding loose connective tissue form may lead in time to irreversible contracture. Tendons and muscles, when maintained in a single position for extend periods also become structurally altered and can shorten permanently. It is not known whether or not the relative inactivity characteristic of many elderly patients can cause similar permanent changes in range of motion with time (Frankel et al., 2006).

The importance of flexibility relative to one's fitness level increases with age. Loss of flexibility impairs most functions needed for good mobility, including bending, stooping, lifting, reaching, walking, and stair climbing. Maintaining lower-body flexibility is also important in low back pain prevention, musculoskeletal injuries, gait abnormalities, and in reducing risk of falling. A reduced range of motion in the shoulder girdle causes significant disability in as much as 30% of the healthy population over 65 years. Lower- and upper-body flexibility, both of which decline with the age but can be improved through exercise, are important aspects of functional fitness for older adults (Rikli & Jones, 2001).

1.3.4 Agility/dynamic balance

Balance is a complex trait and relies on the collective integrity of multiple peripheral and central nervous system components. These include Golgi organs, Ruffini corpuscles, muscle spindles, large myelinated proprioceptive nerve fibers, the posterior spinal cord columns, the medial lemniscus and cerebellum, and the vestibular and visual systems. Together these may be thought of as a 'postural control system', with multiple redundant systems being employed to keep body upright. Deterioration in one or more aspects of the postural control system may occur naturally with age. Consequently, falls are the leading cause of accidental death in older persons. When inactivity declines greatly with the length of a hospital stay or period of an extreme immobility, general deconditioning may adversely affect balance. Additionally, vascular disease, diabetes, excessive alcohol use, medications, and nutritional deficiencies may cause damage to peripheral nerves carrying proprioceptive information. Finally, Parkinson's disease and other common neurologic disorders have been shown to negatively affect balance (Frankel et al., 2006).

Combined agility (involving speed and coordination) and dynamic balance (maintaining postural stability while moving) is important for a number of common mobility tasks that require quick maneuvering such as getting on and off the bus in a timely manner, moving out of the way to avoid getting hit by car or other object, or getting up quickly to answer a phone call, go to the bathroom, or tend to do something in the kitchen (Rikli & Jones, 2001).

1.4 Rationale for functional fitness evaluation

Based on the previous text much of the usual age-related decline in physical ability is preventable and even reversible through proper attention to one's fitness levels and physical activity. Especially important is the early detection in physical weaknesses and appropriate changes in physical activity habits. Until recently, however, most tests to evaluate physical fitness were developed either for young people (resulting in tests that are inappropriate, unsafe, and too demanding to be completed by many older adults) or for more frail elderly to determine the amount of care or assistance needed with activities of daily living. Tests appropriate for frail individuals are too easy and not sufficiently challenging to evaluate fitness in healthier older adults (Rikli & Jones, 2001). The primary goal is to monitor a large population of older adults so that evolving weakness might be identified and treated before resulting in impairment leading to limitations in functional behavior. To create effective exercise interventions the knowledge of an actual functional fitness status is essential. The most important reasons for assessing physical fitness are summarized bellowed and in detail described in the chapter 3.

- Detection of physical weaknesses
- Identifying risk factors
- Planning and evaluating programs
- Educating and setting goals
- Motivating clients

1.5 Statement of the problem and general aims

Functional independence is a primary contributor to quality of life. Technical advances and medical expertise have prolonged life for persons experiencing traumatic injury, stroke, heart failure and chronic conditions. As a result, more people are living longer with the possibility of becoming disabled. Accurate assessment of physical function is important for predicting risk factors for functional dependence, institutional discharge planning, documenting intervention strategies, and medical reimbursement. Optimal physical performance results in the coordinated integration of the cardiovascular and neuromuscular systems into efficient movements which are mediated by psychosocial factors (e.g. motivation, depression, and confidence).

Assessing the physical fitness in older adults is challenging because many factors make this population difficult to test. Factors such as pain, impaired condition, and changes in medical condition and medications increase both within-subject and error variance in measurement. However, the assessment of functional fitness is extremely important in older adults. How and what to assess depends on several factors including objectives of the testing, needs of the participant, age, and existing fitness status. Currently, there are many different standardized tests and test batteries to measure components of physical fitness. Individual tests have been summarized by Carr and her colleagues (Carr, Emes, & Rogerson, 2003) and test batteries which will be summarized and described in the next chapter.

In the Czech Republic there is a lack of relevant information about the older population's fitness functioning measured by performance tests. Recently, one study that used performance tests among older Czech adults been published (Mahrova, Bunc, & Fischerova, 2006). Also Dr. Chytrackova and Dr. Stilec both from the Faculty of Physical Education and Sports, Charles University, Prague have done somewhat extensive work with older adults including among other testing but their result have not yet been published in peer reviewed journal. This particular study has been conducted to fill the gap and to gain an experience with performance testing of older adults in the Czech Republic. The overall goal was to describe in detail a fitness status of Czech older adults and to provide pilot data essential for the future research conducted to develop normative standards for Czech older population.

Performance tests or test batteries developed all around the world, as well as in the Czech Republic, for a younger population are not usable in old age. There is an increasing number of special assessment instruments developed for older adults, but all of them were developed out of the Czech Republic, mostly in the United States. This fact motivated us to transfer an already developed and valid instrument standardized to measure fitness in older adults and test it for the use in older Czech population. The transfer of already standardized tests ensures crucial advantages such as comparability across studies or nations and the ability to accumulate findings. Also a development of new tool is very time consuming and requires a lot of effort from both researchers and participants.

Therefore, the first general aim of this study was to summarize and organize existing instruments developed to measurer fitness among older adults. This extensive review was essential for further steps in this study. Additionally, it may be used as a starting point by anybody who is interested in measuring fitness among older adults either for research or clinical purposes in the future. The priority was to find an instrument which will be well standardized, will measure all functional fitness components, will suffer from minimal ceiling and floor effects, will be sensitive enough to detect changes due to aging itself or due to intervention, and will be accessible for most of the older adults.

The second general aim was to apply the most sufficient battery and to provide the detail information about the functional fitness status of older Czech adults. To make sure that it will be accepted by the majority of older adults (both lower and higher functioning) participants were recruited from two completely different backgrounds creating two sub-samples. The first sub-sample consisted of older adults living independently in the community. Those individuals also regularly participated in leisure activities and they were considered as higher functioning. This sub-sample was also used as pilot study. The second sub-sample consisted of residents of Residential Care Facilities and was considered as lower functioning.

The last general aim will address a question about the structure of the functional fitness. As was documented earlier, functional fitness is composed of several components. These components are considered to contribute to the overall construct and all together should shape the overall score. The functional fitness is an example of the latent construct. Latent constructs are very common in behavioral research and theirs measurement requires special attention and proper methodology. The crucial aspect is to verify the unidimensionality of a measured construct which in other words means to verify that all the components are measuring a single underlying trait. Further, a contribution of individual fitness components to the overall construct was examined. It is logical to expect that some components are more important for the overall construct that other ones. The last aim will not only address the structure of functional fitness and the contribution of relevant components, but it will also provide guidelines for the theoretical and statistical processes that are required to accomplish similar attempt in different areas of behavioral research and beyond.

Summary of general aims of the dissertation

General aim 1: to review, summarize, and organize the most commonly used standardized instruments developed to assess fitness status among older adults over 60 years of age.

General aim 2: to apply the most sufficient instrument measuring functional fitness to higher functioning older adults (pilot testing) and to lower functioning older adults in order to capture a wide range of age and ability levels.

General aim 3: to examine the structure of functional fitness and evaluate the contribution of individual component to the overall construct.

1.6 Dissertation organization

This dissertation is divided into three independent studies each of which was conducted to address above mentioned general aims. Each aim represents a step which was essential to the following one so even though it may look like three independent studies they are actually closely interconnected. The second study builds on results of the first study and the third study extends results of the second study. The first study (general aim 1 – chapter 2) reviewed the available standardized instruments developed to measure fitness among older adults. Since the area of fitness measurement is quite large and not very well organized, this section was focused on the organization of the tests and test batteries into categories based on different methods of measuring. The long term goal of this study was to make it easier for future researches or clinicians to find what they are looking for in the area of measurement fitness in an aged population. Based on the review, the second study (general aim 2 – chapter 3) applied the most sufficient instrument which fulfilled previously set up requirements to measure a population of older Czech adults. The biggest concern was to be able to sensitively test the majority of older adults and provide pilot data for eventual future research conducted to develop normative tables. The results of chapter 3 also served as baseline data for the last study (general aim 3 – chapter 4) which addressed more theoretical issues. The goal was to develop a summary composite measurement score of Overall Functional Fitness. The relatively new approach based on structural equation modeling was applied. This last study answered two specific questions: what is functional fitness and how the overall score can be accurately estimated? From the long term point of view, the process described in this chapter may be used as a model for similar composite measurement score development in any other area of behavioral research and elsewhere.

1.7 Overview of steps of dissertation process

1. An in depth literature review was made and the problem was identified
2. Specific hypothesis-driven aims were described and main instrument was selected
3. Detailed dissertation structure was proposed and approved
4. The instrument was tested in the pilot study
5. Research sample was identified
6. Data were collected and analyzed
7. The whole process was described in this dissertation

CHAPTER 2 – Overview of available instruments evaluating fitness

2.1 Instruments appropriate for older adults

Fitness assessment is essential for maintaining independency and self-supported life in older adults. To successfully prevent the decline of related components we need to know the actual status of those components. There are many different methods each of which has certain advantages and disadvantages.

Evaluation of fitness or physical functioning plays a valuable role in clinical geriatrics as well as in aging research. In the 90's, physical functioning had generally been assessed through self-reports or interviewer-reports. An important addition to this type of subjective assessment is the use of objective performance-based measures of physical function, which have become very popular among clinicians and researchers during last few decades.

Despite the importance of physical functioning, a gold standard measure for this construct does not exist. Self-reported (either self or interviewer administered) and performance instruments have been used in clinical and research settings for decades. However, the phrasing and responses to self-reported instruments have not been uniform. For example, scales may use different reference periods (e.g. 'during last month' vs. 'yesterday') and response items (e.g. 'how much difficulty in performing task' vs. 'how much has your health limited you in performing task'). Some scales classify use of assistive devices or partial assistance as independent, whereas others do not. Even minor differences in wording of response items and threshold for classifying dependency can have substantial effects on prevalence rates of disabilities. Concerns about reproducibility, ability to capture the spectrum of disability, precision, and sensitivity to change of self-report scales have led to the development of performance based instruments. These instruments measure dimensions of physical functioning (e.g. strength, endurance, or balance) or functional tasks (real or simulated) quantitatively, often by timing or by applying objective scoring rules (e.g. counting the number of repetitions, flights of stairs climbed, or measuring time).

Although the clinical and research communities of geriatrics have accepted the importance of measuring physical function, the methods by which this construct is measured have not been yet uniform. Earlier discussed methods have distinct advantages and disadvantages. The results of the study conducted by Reuben and his colleagues (1995) indicated weak to moderate associations between instruments that are designed to measure physical functioning by different methods, suggesting that these instruments are not measuring exactly the same thing (Reuben, Valle, Hays, & Siu, 1995). In another study, it has been documented that older people experiencing decline may fail to report the early changes in physical function (Fried, Bandeen-Roche, Chaves, & Johnson, 2000). This finding inspired Brach et al. (2002) to conduct the study which would compare the descriptions of physical function in community-dwelling older women obtained by performance-based and self-report measures. One hundred seventy community-dwelling women with a mean age of 74.3 years (SD = 4.3, range = 56.6-83.6) completed the activities of daily living, instrumental activities of daily living and social activity sections of the Functional Status Questionnaire (Ware & Sherbourne, 1992). They also completed performance-based measures of gait speed and the 7-item Physical Performance Test (M. E. Cress, Petrella, Moore, & Schenkman, 2005). The majority of the women scored at the ceiling for the self-report measures of function, whereas only 7% scored at the ceiling for the Physical Performance Test and just 30% scored at the ceiling for gait speed. In conclusion, this, performance-based measures identified more limitations in physical function than did self-report measures in this sample of community-dwelling older women (Brach, VanSwearingen, Newman, & Kriska, 2002).

In summary, subjective and objective based measures both capture unique and important information about physical functioning. Because subjective instruments try to assess limitations leading to the disability as well as routine daily activities, when older people become sick bedridden, their answers may not reflect true ability at the time of assessment. In contrast, objective instrument only simulate a given activity, they do not reflect adaptations older people make to facilitate routine day-to-day performance of the activity. Investigators have suggested that the choice of the measure should be determined by research objectives and study population. All available methods of physical functioning are organized in the following figure 2.1.

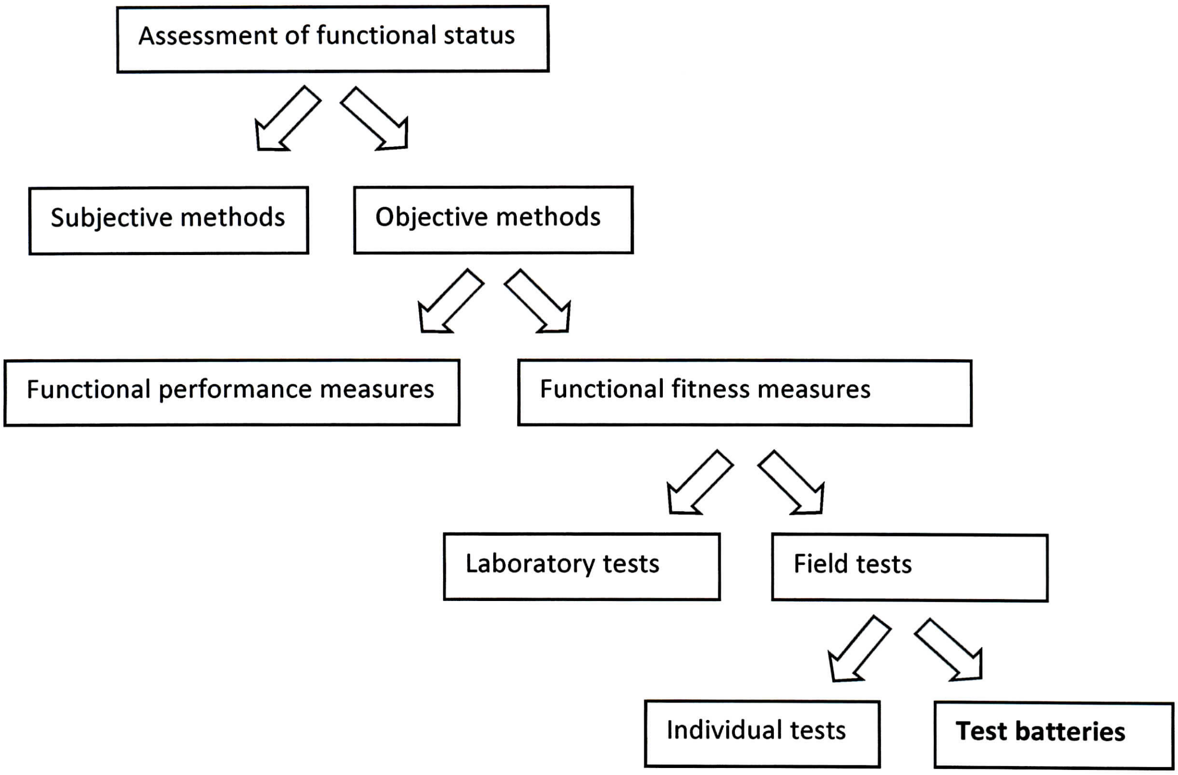


Figure 2.1 Diagram illustrating the tree of available measurement methods

2.2 Available methods of testing and the most commonly used instruments

In this section, the most commonly used examples of each evaluative method as can be seen in figure 2.1 will be addressed. The attention will be focused on objective measures rather than subjective ones.

2.2.1 Subjective measures: self-report and interview techniques

Subjective methods are routinely used to determine the capacity of very old adults. Interviewer simply asks participants or their spouses, relatives, or caretakers and notes those tasks they can accomplish. Another technique to provide an ordered scale might be use of for example Likert scales on which the tested person, or someone else, indicates his or her level of ability. The most commonly used examples of batteries are as follows:

- Index of ADL (Katz, Ford, Moskowitz, Jackson, & Jaffe, 1963)/the Barthel Index (Mahoney & Barthel, 1965)
- Lawton IADL scale – IADL (Lawton & Brody, 1969)
- Functional Independence Measure - FIM (Keith, Granger, Hamilton, & Sherwin, 1987)
- Functional Assessment Measure - FAM (Hall, 1997)
- Physical Disability Index - PDI (Gerety et al., 1993)
- Physical functioning of Short Form 36 questionnaire - SF-36 (Ware & Sherbourne, 1992)
- Advanced Activities of Daily Living - AADL (Reuben, Laliberte, Hiris, & Mor, 1990)

2.2.2 Objective measures: quantification of observed performance

During the past three decades, physical clinicians and researchers have struggled to determine the most appropriate methods to evaluate the ability of elder individuals to maintain their independence in activities of daily living in order to get an objective measure of their functional status. Given that this level of autonomy relies on the effective combination of several physical capacities: endurance; strength; flexibility, and dynamic balance/agility. Some batteries in addition include reaction time, coordination, or speed. Standardized methods for the evaluation of physical performance and functional ability have been developed. In an attempt to enhance the ability to quantify the functional status of the elderly, direct physical performance measures defined as a series of tasks that the individual must perform in a standardized manner using a priori criteria (Branch, Katz, Kniepmann, & Papsidero, 1984; Guralnik, Simonsick et al., 1994) have been developed and gathered under the structure of **functional performance measures** and **functional fitness measures**. Functional performance measures incorporate actual physical skills that are performed on daily basis while functional fitness measures were designed to test basic attributes or capacities thought to underlie daily tasks. Performance is usually evaluated by standardized manners using predetermined criteria which may include counting of repetition or timing of the activity as appropriate. This chapter reviews the most relevant performance batteries that have been specifically designed to assess the functional status of older adults in order to provide knowledge on selection, administration, and interpretation of assessment tools.

2.2.2.1 Functional performance measures

The extent to which an individual can live independently depends largely on his or her ability to perform daily functional tasks known as ADLs. ADLs are the tasks that define an individual's daily functional competence: basic (self-care, hygiene, etc.) and instrumental (household, shopping, etc.). Batteries reviewed in this section include tests focused on the ability to reproduce complex, real-life tasks rather than on specific physiologic abilities. Thus, these measures are closer to the concept of disability than are the tests of more basic abilities. Individuals perform the some kind of daily task and their performance is evaluated in systematic way using mostly continuous scales or predetermined criteria. The most commonly used examples of batteries that fall to this category are in detail described bellowed.

- Physical Performance Test battery - PPT (Reuben & Siu, 1990)
- Physical Performance and Mobility Examination battery - PPME (Winograd et al., 1994)
- Continuous Scale Physical Functional Performance battery - CS-PFP (M. E. Cress et al., 1996)
- Physical Functional Performance 10 - PFP-10 (M. E. Cress et al., 2005)

Physical Performance Test (PPT) (Reuben & Siu, 1990) was created with the idea of obtaining an objective quantifiable measure of functional capabilities. The PPT assesses multiple components of the physical function through the performance of different daily life activities of various degrees of difficulty.

This test battery assesses both lower and upper extremity function. It consists of either nine or seven items version, each scored from 0 – 4, for a total range of 0 – 36 or 0 – 26 in the short version. By including items that extend across a broad range of difficulty (from the easiest to the most difficult), the PPT provides measures of physical function for diverse group of elderly, ranging from those who are dependent in ADLs and those who are fully independent in IADLs but may demonstrate impairment in both groups on an item such as stair climbing. Scoring was based on the time it took to complete each task. Per standardized protocols, time was converted to a 0-4 scale with a higher score representing better performance. Previous research has determined the PPT to be a reliable test and a valid measure of physical performance for community dwelling older adults over the age of 75 years. Low effort tests are: writing a sentence, simulated eating, and turning 360°; medium effort tests are: lifting a book, putting on and removing jacket, picking up a penny from the floor, and walking 50 feet; and high effort test is: climbing stairs.

Physical Performance and Mobility Examination (PPME) (Winograd et al., 1994) was developed to fill the need for a performance measure of physical functioning and mobility appropriate for hospitalized and frail elders. The PPME battery has been designed to screen gross level of function and to detect clinically relevant changes in mobility. The PPME is a reliable and valid performance-based instrument measuring physical functioning and mobility in hospitalized and frail elderly. Tasks were selected so that they can be safely and reliably administered at the bedside, office, or home by nonprofessionals after brief training. The scoring procedure is to yield an adequate distribution of scores (avoiding floor or ceiling effects for frail elderly), and produce the scales that are sensitive to clinically relevant changes. Six mobility tasks of this battery integral to daily life include: bed mobility, transfer skills, multiple stands from chair, standing balance, step-up, and ambulation. Two scoring schemas were developed for each task: 1) dichotomous pass-fail and 2) 3-level high pass, low pass, and fail. A summary scale was developed for each method of scoring.

Continuous Scale Physical Functional Performance battery (CS-PFP) (M. E. Cress et al., 1996) is a unique instrument designed to provide a comprehensive, in-depth measure of physical function that reflects abilities in several separate physical domains. The CS-PFP is based on ordinary activities of daily life, performed at maximal effort within the bounds of safety and comfort. The CS-PFP consist of battery of sixteen everyday tasks, ranging from easy to demanding, that sample the physical domains of upper and lower body strength, upper body flexibility, balance, coordination, and endurance. Each task is scored 0 to 100 based on an empirically derived range established from data gathered on older adults with a broad range of individual functional abilities. The test yields a total score up to 100 that is the average of five separate physical domain scores. Testing requires standard laboratory conditions and approximately one hour to administer. All tasks are quantified by time, distance, or weight. This test is a valid, reliable, and sensitive measure of physical function. It is applicable to a wide range of function levels and has minimal floor and ceiling effects. The test may be used to evaluate, discriminate, and predict physical functional performance for both research and clinical purposes. Low effort tests are: carry a weighted pan a distance of 1 m, pouring water from a jug into a cup, donning and removing a jacket, place a sponge on and remove it from an adjustable shelve; medium effort tests are: floor sweeping with a broom and dustpan, transfer clothes from washer to dryer and transfer the clothes from dryer to basket, making a bed, vacuuming, place a strap over a shoe, open and pass through a fire door, pick up four scarves from the floor; and hard effort tests are: carry a weighted bag up and down simulated bus stop, carry a groceries 70m, sit and stand up from the floor, climb stairs, and 6-minute walk.

Physical Functional Performance 10 (PFP) (M. E. Cress et al., 2005) was developed to validate a short version of CS-PFP that requires less space and equipment. The following criteria were used to determine which items from the original CS-PFP to include in the shorter version of the test:

1. All items were retained to preserve the integrity of that domain, even if they did not discriminate across the three groups.

2. Tasks that did not discriminate among the three groups were removed in order to eliminate ceiling or floor effects on individual tasks.

3. When two items contained similar components, the item that provided the most information was retained. For other items, the item that had the least burden of setup and administration was retained.

The test that remained in the short version are: carry a weighted pan a distance of 1 m, donning and removing a jacket, place a sponge on and remove it from an adjustable shelve, floor sweeping with a broom and dustpan, transfer clothes from washer to dryer and transfer the clothes from dryer to basket, pick up a four scarves from the floor, carry a groceries 70m, sit and stand up from the floor, climb stairs, and 6-minute walk.

2.2.2.2 Functional fitness measures

As in previous test batteries, participants perform an actual task but the task does not replicate daily activities. It is designed to measure the capacity which is necessary to independently perform daily activities. In other words, those tests were developed to evaluate functional fitness and its components rather that functional performance. Most functional fitness tests were firstly developed for younger subject. But most of tests for younger population are not sufficient for use among older adults. Therefore it was necessary to modify existing tests or test batteries or to design completely new ones. Functional fitness testing may be further divided into sub-groups based on where the testing is performed and if the single or multiple components are tested. The sub-groups are as follows:

- Laboratory tests
- Field tests
    - Individual tests
    - Test batteries

Laboratory tests

Laboratory test are usually considered to be 'gold standards' for measuring the physiological parameters such as aerobic endurance, muscle strength, flexibility or agility. Those tests are taken in special laboratories furnished by special equipment such as treadmill, ergo meter, dynamometry, goniometry, or special stability platforms and related technologies. Laboratory testing is expensive, time consuming, and sometimes not appropriate for the majority of older adults. For example the gold standard for aerobic capacity/fitness is the maximum oxygen uptake – VO_2max. It is based on the stressing body to maximal exhaustion which is not very often appropriate for older adults therefore the sub-maximal tests which are less demanding can be used. Laboratory tests are mostly used for testing either elite older adults or younger older adults. A few of the most common examples are described bellowed.

Bruce Protocol Stress Test (Bruce, Kusumi, & Hosmer, 1973; Spirduso, 2005) is a commonly used treadmill exercise stress test. It was developed as a clinical test to evaluate patients with coronary heart disease, though it may also be used to estimate cardiovascular fitness. The test starts at 2.74 km/hr at the gradient of 10 %. At three minutes intervals the incline of the treadmill increases by 2%, and the speed increases as shown in the table 2.1.

Table 2.1 Protocol for Bruce treadmill stress test

Stage	Speed (km/h)	Gradient
1	2.74	10
2	4.02	12
3	5.47	14
4	6.76	16
5	8.05	18
6	8.85	20
7	9.65	22
8	10.46	24
9	11.26	26

Another example may be the **Modified Bruce Protocol Stress Test**, which starts at a lower workload than the standard test, and is typically used for older adults or even young but low functioning patients. The first two stages of the Modified Bruce Protocol are performed at 2.74 km/hr at the 0% gradient and 2.74km/hr at the 5% gradient. The third stage corresponds to the first stage of the standard Bruce Test Protocol as described above. The test score is the time taken on the test in minutes. Another example may be **Balke Treadmill Test** (American College of Sports Medicine, 2000) which was developed as a clinical test to determine peak oxygen uptake in cardiac patients. The **Astrand-Rhyming Test** (American College of Sports Medicine, 2000) is an example of sub-maximal test performed on ergometer at constant workload for seven minutes. But here, the test score would be influenced by the variability in maximum heart rate in individuals. It would underestimate the fitness of those with high maximum heart rate, and overestimate fitness with advancing age (as max HR reduces with age). As it is performed on the cycle ergo meter, it would favor cyclists.

Field tests

Individual tests (just examples)

- Muscular strength: 30-second chair stand test, arm curl test, grip strength test, leg-extension strength test, leg muscular endurance test, timed 10 chair stands, timed 5 chair stand test, stair climbing.
- Aerobic endurance: half mile or mile walking tests, self-paced 2-minute step test, 3-minute step test, 5-minute walk test, 6-minute walk test, 9-minute walk test, the Cooper test.
- Flexibility and range of motion: standard sit-and-reach test, chair sit-and-reach test, back scratch test, Appley scratch test of internal rotation, circumduction test.
- Agility or balance: 8-foot up-and-go test, get up-and-go test, Romberg test and sharpened Romberg test, balance board test, Berg's balance scale, test of standing balance, one leg stand.

Test batteries

Functional fitness has been defined as having the physiologic capacity to perform normal everyday activities safely and independently without undue fatigue (Rikli & Jones, 2001), so batteries must assess the physiologic attributes that support the behavioral functions necessary to perform ADLs. They are usually consisted of different combinations of individual tests mentioned earlier. Functional fitness is typically assessed using batteries that include measurements of aerobic capacity, muscular strength and endurance, flexibility, and agility. In some cases the motor performance might be added. This most often includes components such as speed, coordination, or static balance. The most common examples of batteries that fall to this category are in detail described below.

- American Alliance for Health, Physical Education, Recreation, and Dance battery - AAHPERD (Osness, Adrian, & Clark, 1996)
- Short Physical Performance Battery - SPPB (Guralnik, Simonsick et al., 1994)
- MacArthur battery (Guralnik, Seeman, Tinetti, Nevitt, & Berkman, 1994)
- Functional Fitness Battery - FFB (Netz & Argot, 1997)
- Senior Fitness Test battery - SFT (Rikli & Jones, 2001)
- The Groningen Fitness test for the Elderly - GFE (Lemmink, Han, de Greef, Rispens, & Stevens, 2001)
- Health ABC battery (Brach, Simonsick, Kritchevsky, Yaffe, & Newman, 2004)

AAHPERD battery (Osness et al., 1996) was one of the first attempts to create a specific battery to measure the physical fitness of the elderly. It was carried out by the American Alliance for Health, Physical Education, Recreation and Dance (AAHPERD). This battery, also known as Functional Fitness Assessment Battery, has become one of the most popular batteries and one of the most useful instruments. However, certain weak points have been detected, such as the absence of some lower-body muscle function tests or the fact that some of the exercises (flexibility and aerobic endurance) may be difficult to perform for many older participants. All in all, the AAHPERD is a battery which can be easily administered due to its low cost and its minimal space and equipment requirements. It also includes a large number of reference parameters and a user's manual to apply it. Only the learning effect needs to be controlled, especially as far as the agility, flexibility, and coordination exercises are concerned. It has been considered as a useful instrument to evaluate the underlying physical parameters associated with daily activities. The following tests are included: arm curl test, timed 880-yard walk, standard sit-and-reach test, walking course (participants started from seated position and were asked to rise from the chair, walk around a cone located 6 feet to the left of and 5 feet behind the chair, return to the seat and sit down, stand again, walk around cone 6 feet to the right of and 5 feet behind the chair, return to the seat and sit down, and the repeat the entire procedure), and soda pop test. In addition, the ponderal index, a height-weight ratio, was computed as an index of body composition.

Short Physical Performance Battery (SPPB) (Guralnik, Simonsick et al., 1994) known also as NIA (National Institute on Aging), or EPESE (Established Population for the Epidemiologic Studies of the Elderly) battery. It was derived from the adaptation of different functional tests created during the 1980s with the objective of being administered by one single person, in any home, regardless of any spatial constraints. The detailed instructions may be downloaded from the following web site:
http://www.grc.nia.nih.gov/branches/ledb/sppb/index.htm
The SPPB battery is focused on assessing the lower extremity function and is able to classify a large number of elderly people across a broad spectrum of functional status. This is one of the most widely used batteries in longitudinal studies seeking to evaluate elderly people. This battery is characterized by the short period of time involved in its performance (10 - 15 min) and by the fact that it predicts mobility disability and activities of daily living disability independently, mainly through the assessment of strength, balance, and gait speed. The following tests are included: 5 chair stands, tandem, semi-tandem and side-by-side stands, and 8-foot walking course.

In addition, categories of performance were created for each set of performance measure to permit analysis that included those unable to perform a task. For 8-foot walk and repeated chair stands, those who could not complete the task were assigned a score of 0. Those completing the task were assigned scores from 1 to 4 corresponding quartiles of time needed perform a task, with the fastest time scored as 4. The three test of standing balance were considered as hierarchical in difficulty in assigning a single score of 0 to 4 for standing balance. A summary performance scale was created by summing the category scores for walking, chair stands, and balance tests and ranged from 0 to 12.

MacArthur battery (Guralnik, Seeman et al., 1994) was created for the use in the MacArthur Study of Successful Aging investigating the factors that influence physical and cognitive functioning among relatively high functioning volunteers between the ages of 70 and 79, with the main objective of identifying the key factors that seem to contribute to healthy aging. Tests included in the battery represent several major domains of physical performance that were derived from previous studies. Because of that, these measures have generally good reliability. While the battery presents performance measures of functioning as true measures of physical health status in nondisabled old persons, it should be noted that some important domains, such as upper extremity strength and shoulder range of motion, have not been considered. The following tests are included: hand grip strength, balance, 10-ft walk, tapping a foot, to sing a name, switching back and forth between two circles 1ft apart while in a seated position.

Functional Fitness Battery (Netz & Argot, 1997) was developed to provide a field test to assess various components of daily activities. This battery consists of eight subtest components which try to reproduce daily activities, measuring the fitness level at the same time. Three of these components have been taken from the AAHPERD battery, while the rest of them are completely new. This battery does not require special equipment, it is a low-cost battery easy to perform, and it does not involve too much time (50 people can be tested in 3h). Its strong point is that it is easy to administer and does not need a doctor's permission. However, it is only useful to assess independent individuals (which is a limitation) and large populations. It might not be a good choice if the purpose of the study is to carry out clinical studies or pre-test and post-tests over a short period of time.

Senior Fitness Test (SFT) (Rikli & Jones, 2001) was developed to assess key physiologic parameters. It is also known as Fullerton Fitness Test. This battery focuses on the evaluation of those physical abilities which allow the functional independence of the older adults. The SFT is relatively easy to perform, the exercises are safe, it has almost no ceiling and floor effects, and there are "normative scores" for each exercise, which makes of it a very useful battery to assess the functional fitness. Besides, if it is organized as a circuit, it is possible to evaluate up to twenty four people in 90 min. It is worth mentioning that in spite of the fact that the construct validity of the SFT has been confirmed, some kind of learning effect has been detected, and therefore, one or two previous practice sessions are advisable prior to the final assessment session. Lastly, it must be taken into consideration that the SFT has been created and validated upon the score of voluntary elderly people, with ambulatory independence and generally active; consequently, the extrapolation of the normative scores should be done with caution. The following tests are included: chair stand test, arm curl test, 6-minute walk test or 2-minute step test, chair sit-and-reach test, back scratch test, and 8-foot up-and-go test.

The Groningen Fitness test for the Elderly (GFE) (Lemmink et al., 2001) is a field based fitness assessment designed to research the interrelationship between physical fitness, physical activity, health, and daily functioning. This battery, compared to others, includes manual dexterity and reaction time tests and it is combined with a questionnaire to assess the subjective self-evaluation of health. The reliability, inter-rater, and internal consistency of the GFE have been demonstrated, which makes of it a very useful tool to measure basic fitness components such as strength, endurance, and coordination. Although "passing" the GFE does not take long (each test takes 4 minutes and the endurance test takes 15 minutes) and its items are simple to perform and easily transportable, it does require specific equipment. In addition to this, it should be noted that the circumduction test lacks objectivity, that including a suitable warming up before the sit and reach test has been suggested, and that some previous practice before the block-transfer exercise would be advisable in order to avoid the learning effect. Furthermore, the endurance exercise (walking test) may not be selective enough, given the fact that some people are able to finish it without reaching their maximum level of effort. Lastly, it must be stressed that because this battery consists of simple exercises and it is easy to administer, it is used to assess the fitness level of sedentary populations and of people affected by different pathologies. The following tests are included: grip strength, walking endurance test with increasing speed, sit and reach test, circumduction test, balance board test, gait speed, block transfer, and single reaction time.

Health ABC (Brach et al., 2004) was developed for use in the Health Aging and Body Composition Study (Health ABC) which was "a prospective investigation of interrelationships between health conditions, body composition, social behavioral factors, and change in physical function". To measure a wider range of function in this population, the SPPB battery was expanded to create the Health ABC performance battery. Because of that, the hold times on the standing balance used in SPPB battery were increased to 30s, and two additional balance tests were added. Besides, walking endurance is usually assessed by means of the 400-m walk test. Although this battery includes a large reference database to compare scores, it is important to point out that the selective criteria for joining the ABC study are being able to walk a quarter of a mile, climbing up ten steps, or performing basic daily life activities. The following tests are included: chair stands, isokonetic strength, 400-m walk, parallel feet, semi-tandem, and tandem stand, single leg stand, 20-cm narrow walk, and gait speed.

Table 2.2 Overview of functional fitness measures and relevant components included in each battery

Functional fitness component	**AAHPERD battery**	**SPPB battery**	**Mac A. battery**	**FFB battery**	**SFT battery**	**GFT battery**	**ABC battery**
Lower-body muscle strength		**Yes**		**Yes**	**Yes**		**Yes**
Upper-body muscle strength	**Yes**		**Yes**	**Yes**	**Yes**	**Yes**	
Aerobic endurance	**Yes**			**Yes**	**Yes**	**Yes**	**Yes**
Lower-body flexibility	**Yes**			**Yes**	**Yes**	**Yes**	
Upper-body flexibility				**Yes**	**Yes**	**Yes**	
Agility/ dynamic balance	**Yes**		**Yes**	**Yes**	**Yes**		
Static balance		**Yes**		**Yes**		**Yes**	**Yes**
Speed		**Yes**		**Yes**		**Yes**	**Yes**
Coordination	**Yes**		**Yes**			**Yes**	

2.3 Steps of process of instrument selection

Researchers, physical therapists, or fitness instructors should consider several factors before selecting an assessment instrument to measure functional fitness of older adults:

1. What is the purpose of the assessment? Instruments measure a variety of physiologic parameters and functional activities. Instructors should select an assessment instrument based on the reason of the assessment. For instance, if the purpose is to evaluate the effect of an intervention program established to improve functional fitness, all relevant components of functional fitness should be measured and the instrument should be sensitive enough to detect desired changes.

2. What is the actual functional ability level of the participants? Some assessment instruments are specifically designed for use with either frail or healthy older adults resulting in either ceiling or floor effect when inappropriately used. Few are able to assess the wide range of ability levels that most senior fitness instructors encounter. It is desirable to select test items that can be used with a wide range of functional abilities—from the borderline frail to the highly fit. Spirduso (2005) described five levels of functional abilities as follows:

   - Physically dependent: ability to pass only some basic activities of daily living (BADLs). These activities include walking, bathing, dressing, or eating. Clearly if an individual is unable to manage these activities, they need care in either their own home or care facility.

   - Physically frail: capable of instrumental activities of daily living (IADLs) such as light housekeeping, food preparation, and grocery shopping. They are competent in all BADLs, but many have problems with some IADLs.

   - Physically independent: routinely engage in low physical demand activities such as golf, crafts, woodworking, traveling, and driving. They also have the ability to do IADLs.

- Physically fit: capable of doing most hobbies and competent in moderate physical activities. They will also participate in exercise 2-3 times each week for their health and well-being.

- Physically elite: capable to train heavily on daily basis with the goal of participating in a competition. They are a rare and unique population who may complete in marathons, senior Olympics, and master tournaments.

3. Is the assessment instrument reliable and valid for use with the older adult population? This is one of the most important factors to consider when selecting an assessment instrument. Validity refers to whether a test item measures what it is intended to measure; reliability refers to the repeatability of test scores. A reliable test produces similar scores from one trial to another, or one day to the next. Scoring accuracy is an important test characteristic. Instructors should never assume that tests developed and validated for younger age groups will be appropriate for older adults. They should check that the test items being considered have supporting data documenting its reliability and validity based on studies involving older adults.

4. Is the assessment instrument feasible to use? Feasibility refers to the suitability of administering the test items in a given setting. Instructors should consider the following additional questions to determine feasibility:

   - What type of equipment is needed?
   - What is the cost of the equipment?
   - How long does each test item take to administer?
   - How much space is needed?
   - How difficult are the test items to administer and score?
   - Can non-professionals administer the tests?
   - Are the test items safe to give?

In summary, this chapter reviewed all possible methods of fitness evaluation. All instruments were organized and the most commonly used and the most relevant instruments were described in detail. This review was done to provide a logical guide through fitness testing among older adults. It is expected that this review will be helpful not only for purposes of the present dissertation but for anybody to chose the best instrument for any study involving functional status of older adults. This chapter also addressed the process of instrument selection which might be helpful for decision making.

The selection criteria or requirements for this dissertation were established as follows. The first requirement was that the instrument must be standardized for use among older adults. The second requirement was that the instrument must be able to accurately measure all components that are related to daily functioning. This requirement excluded all individual tests and narrowed our attention to test batteries but only to those that measure fitness components, not the daily functioning. Also self-reported instruments were considered as insufficient. The third requirement was that the instrument was developed to be administered in a field setting. The priority was given to quick-and-easy to administer instruments without a need of special equipment and space. The final selection was made from table 2.2 which presents are sufficient functional fitness batteries. Based on our requirements and presented recommendations about how to select the most appropriate instrument, the Senior Fitness Tests appeared to be the best and logic choice for this study.

CHAPTER 3 – Functional fitness among older Czech adults

3.1 Abstract

Background: The ability to measure one's fitness status is essential for both research and clinical purpose. The aim of this study was to apply the Senior Fitness Test (SFT) to two groups of older Czech adults recruited from different backgrounds in order to describe their functional fitness; compare both sub-samples; examine the function of age; and determine what percentage of each studied sub-sample is living at the risk of independency loss.

Methods: Ninety three older adults (> 60 years) were included in the study. Fifteen participants lived independently in the community (9 women and 2 men; mean age 73.7; SD 5.95) and seventy eight older adults were permanent residents in the Residential Care Facilities (RCF) (69 women and 9 men; mean age 81.9; SD 9.31). Functional fitness was assessed by six SFT tests developed to measure major components. Means and standard deviations were performed to describe both sub-samples. Independent-samples *t*-test and one-way ANOVA were used to determine the differences of means between and within sub-samples.

Results: Participants living independently performed significantly better compared to residents from RCFs in the test 1: 30-second chair stand test (t (86) = -5.05, p = .000); test 2: 30-second arm curl test (t (90) = -5.24, p = .000); test 3: 2-minute step test (t (75) = -6.57, p = .000); test 5: back scratch test (t (66) = -3.56, p = .000); and test 6: 8-foot up-and-go test (t (62) = 4.95, p = .001). No statistically significant differences were found between age categories except for test 1: 30-second chair stand test (F (63) = 4.85; p = .011). More than 50% of RCFs' residents were performing at or below the threshold scores associated with the risk of independency loss.

Conclusions: The SFT was well accepted by all participants and it is suitable measure providing valuable information about functional fitness status and may be recommended for the use in the Czech Republic in the future. Development of population norms would be desirable.

3.2 Introduction

Rationale for maintaining functional fitness in old age

The ultimate objective of the successful aging paradigm is to maintain or improve the quality of life among older populations (Morley, 2003; Spirduso, 2005). Although the aging process is currently immutable, one can modify the pathology and mitigate the expression of disease through prevention and treatment. Avoiding or postponing diseases and disabilities, maintaining mental and physical function are the cornerstones of this approach (E. Cress et al., 2008). The importance of physical exercise has been recognized more than two thousand years ago by Hippocrates (450 BC):

> *"All parts of the body which have a function if used in moderation and exercised in labors in which each is accustomed, become thereby healthy, well developed, and age more slowly; but if unused and left idle they become liable to disease, defective in growth, and age quickly."*

The current evidence supporting the advantages of physical activity and exercise interventions into daily routines is vast. Our personal fitness level is important at any stage of life but its importance even increases with the age. As we get older, however, the focus shifts from health promotion to functional independence which is generally referred to as **functional fitness**.

Functional fitness is defined as having capacity to perform normal everyday activities safely and independently, without undue fatigue (Rikli & Jones, 2001). One of the easiest ways to positively influence functional fitness is to pay attention to our physical activity and exercise levels. Older adults, both men and women, can benefit from an active lifestyle. Regular physical activity has the potential to sustain an active lifestyle and maintain independence by increasing ones' functional fitness status. Furthermore, it has been proven that better physically functioning older adults suffer from less chronic diseases (Booth, Laye, Lees, Rector, & Thyfault, 2008); have reduced symptoms of anxiety and depression, and better mood as well as feelings of well-being (Garatachea et al., 2008; Strohle, 2008); rate theirs subjective health better (Jylha, Guralnik, Balfour, & Fried, 2001); have lower risk of mortality (Nocon et al., 2008); and the most importantly, they experienced higher quality of life (Wood et al., 2005).

Functional fitness is a complex construct that includes different components which are essential for maintaining a range of daily activities from the basic such as basic ADL to the most demanding such as traveling or participation in sport. The major components are: muscular strength (upper-body and lower-body); aerobic endurance; flexibility (upper-body and lower-body); and balance (Frankel et al., 2006; Rikli & Jones, 2001; Spirduso, 2005). Improving functional fitness components via targeted physical activity and exercise will enable older adults to adopt more active lifestyles and thus enhance their quality of life. The main aim in creating exercise programs for older adults should be to help those individuals meet their functional performance goals by improving their functional fitness. To be able to successfully accomplish this, we need to know what are the weaknesses and strengths of each individual and what needs to be improved upon.

Importance of functional fitness assessment in older adults

The accurate assessment of components of functional fitness is essential not only to address different needs and ability levels, but also to identify at-risk patients and effectively target exercise intervention programs. The ability to accurately evaluate functional fitness status is essential for both clinical and research purposes. In summary, the main reasons why it is so important to assess the functional fitness in older adults are: program planning and evaluation; identification of at-risk older adults; and goal setting and motivation of older adults.

Program planning and evaluation

To plan safe and effective exercise or physical activity programs for older adults, it is important to know as much as possible about the participant's health and actual fitness status, current physical activity level, activities likes and dislikes, and personal goals. A comprehensive functional fitness test provides specific information regarding a participant's physical strength and weaknesses associated with functional tasks and activity goals important for everyday living. This information is necessary to design individualized, targeted physical activity or exercise programs for older adults. Also, baseline measures repeated at multiple intervals during the program provide critical data to track the progress of participants, to make program adjustments, to provide personalized feedback, and to evaluate program effectiveness.

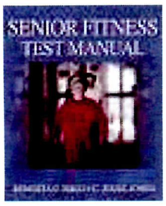
Identification of at-risk participants

Many independent older adults, often due to their sedentary lifestyles, function dangerously close to their maximum ability level during normal activities. Climbing stairs or getting out of a chair requires near maximum effort for many older individuals (Evans, 1995). Early identification of physical decline and appropriate interventions may help to prevent functional impairments. One goal of fitness practitioners should be to identify at-risk participants, and to provide a targeted intervention program or make appropriate medical referrals for a complete diagnosis, treatment and maintenance plan.

Goal setting and motivation of participants

Assessing the functional fitness levels of participants is a precursor to helping them set worthwhile short- and long-term personal goals. To facilitate goal setting, it is especially helpful to relate the purpose and results of assessments to the types of daily tasks and activities older adults hope to continue to do or want to improve. For example, upper body strength is important for performing household and other activities that involve lifting and carrying groceries, suitcases, or grandchildren. So a short-term goal may be to improve upper body strength by 20%, while a long-term goal may be to be able to do yard work or pick up a grandchild. The periodic assessment and monitoring of performance motivates older adults. It also encourages their progress and exercise compliance. Testing also encourages many people to pay more attention to their fitness and physical activity level.

There are many available options to evaluate functional fitness from the most demanding performance tests done in special laboratory settings to self-report assessments (for more details see previous chapter). Each method has certain advantages and disadvantages. From the available selection of standardized test batteries, the Senior Fitness Test (Rikli & Jones, 2001) happened to be the most appropriate for our study. This test battery satisfied our requirements for well standardized, easy to use in field settings instrument developed to evaluate all major functional fitness components among older adults. Additionally, the SFT has the ability to capture functional fitness among wide range of older adults (both lower and higher functioning older adults) with minimum floor or ceiling effects.



The Senior Fitness Test (SFT) was developed to assess the components of functional fitness, specifically: strength – upper-body and lower-body, endurance, balance, and flexibility – upper-body and lower-body. The test battery includes six seven tests assessing six earlier mentioned components (two tests measure the same component - aerobic endurance): 30-second chair stand test measuring **lower-body strength**, 30-second arm curl test measuring **upper-body strength**, 2-minute step test or 6-minute walk test measuring **aerobic endurance**, chair sit-and-reach test measuring **lower-body flexibility**, back-scratch test measuring **upper-body flexibility** and 8-foot up-and-go test measuring **dynamic balance or agility**. Each test is scored separately. For those studies conducted within U.S., results can be compared with normative standards that were developed from a national study of 7000 independently living men and women aged between 60-94 years. Additionally, results may be compared with threshold scores indicating the risk of independency loss. Advantages of SFT are summarized as follows:

- The SFT is comprehensive. The tests reflect a cross section of all major fitness components associated with independent functioning in late years, whereas other test batteries for older adults focus only on selected aspects of fitness.
- The SFT provides continuous-scale measures which helps to sensitively measure a wide range of the population. A common limitation in other test batteries as that some items tend to be either too easy or too difficult for a large portion of older adults.
- The SFT is usable in the field setting. The tests have minimal equipment and space requirements and the entire battery can be administered in most clinical and community settings, as well as in peoples' homes.
- The SFT is very well standardized. Rikli and Jones have shown that for community dwelling adults over 60 years of age, this test has content validity established through literature review and expert opinion, criterion validity correlation coefficient ranging from r = 0.73 to r = 0.83 when comparing each test item with an earlier established criterion measure, and high test retest reliability with correlation coefficients ranging from r = 0.80 to r = 0.98 for the tests.

3.3 Literature review

Application of Senior Fitness Test (studies published up to date in peer review journals)

So far, the SFT has been widely used in many studies (Alexander, Phillips, & Wagner, 2008; Beck, Damkjaer, & Beyer, 2008; Carvalho, Marques, & Mota, 2008; Cyarto, Brown, Marshall, & Trost, 2008; DiBrezzo et al., 2005; Dobek, White, & Gunter, 2006; Garatachea et al., 2008; Mahrova et al., 2006; Thompson, Cobb, & Blackwell, 2007; Toraman, 2005; Toraman & Ayceman, 2005) within and outside of the United States.

The SFT was repeatedly applied to relatively healthy U.S. residents living in the communities. For example, in the study by Cyarto et al. (2008), the SFT was used to compare the effectiveness of home-based and group-based progressive resistance training programs and a group walking program on functional performance in older adults. The participants were residents of a retirement village aged 65 to 96 years (Cyarto et al., 2008). Another example is the study that used the SFT to examine the effects of a simple, low-cost fall prevention exercise program for community dwelling older adults (DiBrezzo et al., 2005). Participants attended a 10-week exercise class that included stretching, strengthening and balance-training exercises. At the completion of the program, significant improvements were observed in tests measuring dynamic balance and agility, lower and upper extremity strength, and upper extremity flexibility. A year later, Dobek et al. (2006) tested the effect of a 10-week exercise program. The purpose of this study was to determine the degree to which a novel training program, based on activities of daily living (ADL), would affect performance of ADLs as well as fitness levels of older adults. Fourteen individuals (mean age 82 years) took part in a 10-week control period followed by a 10-week ADL-based training program. Pre-tests and post-tests included the Physical Performance Test, the Physical Functional Performance-10, and the SFT. The results supported the hypothesis that this novel ADL-based training program facilitated improved not only the performance of ADLs, but also selected measures of functional fitness among older adults. Two years later, the study focused on training followed by detraining in functional fitness components (Carvalho et al., 2008) was conducted. Training induced significant improvements in chair stand test (27.3%), arm curl test (17.4%), chair sit-and-reach test (17.4%), up-and-go test (11%) and back scratch test (14.5%). However, both upper- and lower-body strength and upper- and lower-body flexibility declined significantly after detraining in the exercise group. The results of this study, except for usefulness of SFT, highlighted the negative effects of interrupting. Three years earlier Toraman (Toraman, 2005) used the SFT to test the effects of short (six weeks) and long (fifty two weeks) term detraining in elderly. The authors also determined whether these effects differ according to age. Authors came up with the evidence that even though functional fitness improved during the exercise training period, just short term detraining caused a loss of this improvement.

Performance in all tests reverted to the pre-training values or lower after fifty two weeks of detraining in both groups. The results were confirmed in his next study published in the same year with the additional conclusion that changes in lower extremity flexibility test, timed up-and-go test, and the 6-minute walk test performances in response to six weeks of detraining were affected by age in older adults. In this study, the total of twelve young-old subjects (aged 60-73 years) and nine older subjects (aged 74-86 years) were compared (Toraman & Ayceman, 2005).

Although the SFT has been widely used in United States, it has been recently adopted by other countries. Currently, there is one peer reviewed published article reporting the application of SFT in the Czech Republic (Mahrova et al., 2006). The aim of this study was to choose an acceptable motor test battery for older adults suffering from chronic renal failure. The authors wanted to cover all components of functional fitness needed for performing activities of daily living and to be able to successfully detect the effect of an exercise intervention program designed for twenty three patients. Besides other findings, the Senior Fitness Test was considered an acceptable measurement tool for testing functional fitness among renal dialysis patients. In addition to using the SFT in the Czech Republic, this test battery was also used in Spain. That study investigated the relationship between physical function, well being and physical activity (Garatachea et al., 2008).

SFT is not only acceptable for relatively healthy older adults and patients with chronic renal failure, but also, it was successfully applied in older patients with chronic lung disease (Alexander et al., 2008). The purpose here was to compare the effects of strength training and traditional pulmonary rehabilitation program on functional fitness in older patients with chronic obstructive pulmonary disease. The results showed that the SFT was a sensitive tool for detecting changes in functional fitness and that strength training had a favorable impact on functional fitness in older patients suffering from chronic lung disease.

Despite the fact that the SFT was developed for community dwelling older adults, as documented the test battery was successfully used among older adults suffering from chronic conditions, and also in residents of nursing homes (Beck et al., 2008). The aim of Beck's research was to evaluate the effect of a multifaceted intervention consisting of nutrition, group exercise and oral care on functioning. Functional fitness of participants was assessed by four of the six performance tests from SFT battery (except for flexibility). Because lower levels of performance were expected, a modified version of SFT was used (e.g. use of an arm rests, assistive device if needed such as cone or walker, and lighter hand weights). According to the authors, the performance testing was well tolerated by nursing home residents. Most importantly, the ability of the SFT to detect changes in functional fitness remained even in a low functioning sample.

On the other hand, the SFT was used among elite functioning older adults to examine the effect of a progressive functional training program on golf club head speed and functional fitness in male golfers (mean age 70.7 ± 9.1 [SD] years) (Thompson et al., 2007). The exercise group participated in an 8-week progressive functional training program including flexibility exercises, core stability exercises, balance exercises, and resistance exercises. Pre-tests and post-tests included club head speed of a driver by radar and SFT scores. In summary, this functional training program resulted in significant improvements in club head speed as well as in several components of functional fitness.

In conclusion, the cited articles demonstrate the increase in use of SFT battery. More than half of the articles have been published within last couple years and so far, five articles were published just recently (in 2008). It is also interesting that there is an increased use among lower functioning older adults. Also this instrument outreached the borders of United States. Although the majority of the cited applications of SFT have been applied in community dwelling, relatively healthy older adults, there is an increasing evidence for its use in patients suffering from chronic conditions. Additionally, the SFT with minor adaptation was successfully used in nursing home settings where residents are mostly low functioning and performance tests are rarely used. Vice versa, the SFT was found very useful among very fit – elite older adults. None of the cited studies reported either floor or ceiling effects. The floor effect may occur when too many participants are not able to complete test battery and thus most of them reach the minimum possible scores. A ceiling effect is exactly the opposite situation. It occurs when the test battery is too easy and many participants reach the maximum scores. In conclusion, it was clearly demonstrated that SFT is an effective tool which is able to identity older adults of various functional fitness levels.

3.4 Summary and hypothesis

It has been proven that older adults, both men and women, can benefit from a physical activity. In order to create effective physical activity programs that would positively affect quality of life in later years, we need to know where to focus our attention the most. Therefore, the assessment of all components related to independent living becomes essential. The components of interest are: lower-body and upper-body muscle strength; aerobic endurance; lower-body and upper-body flexibility; and agility. The overall capacity of the organism to perform independently on daily bases is called Functional Fitness. Based on the literature review, the SFT has been identified as an appropriate instrument for testing all components. The SFT battery also fulfilled our criteria therefore it has been selected for the study. The long term goal was to gain an experience with performance testing among older adults living in the Czech Republic. This experience is crucial for future research conducted to develop norms for older Czech population. Findings from this study will provide a baseline data with the detailed information about functional fitness levels of older Czech adults which might be very valuable not only for everybody who works with this population, but also for older adults themselves. Four hypothesis-driven specific aims will be address in this work:

Specific aim 1: to provide the evidence that the Senior Fitness Test is applicable for the majority of older adults living in the Czech Republic.

Specific aim 2: to provide the evidence that the Senior Fitness Test is sensitive enough to detect differences between higher functioning and lower functioning older adults.

Specific aim 3: to provide the evidence that the Senior Fitness Test is sensitive enough to detect differences between age groups.

Specific aim 4: to examine the possibility to successfully identify older adults who are at-risk of independence loss using cutting threshold scores developed in U.S.

Hypothesis:

H1 The Senior Fitness Test is applicable for independently living older adults as well as for residents from Residential Care Facilities.

H2 Older adults living independently in their homes perform significantly better compared to older adults living in Residential Care Facilities in all tests.

H3a All the components of functional fitness decline significantly with age.

H3b The relative age-related decline in functional fitness components is consistent for all tests.

H4a All older adults living in homes will perform above the cutting threshold scores in all tests.

H4b More than 50% of older adults living in Residential Care Facilities will perform at the cutting threshold scores or below it in all tests.

3.5 Methods

3.5.1 Participants

A total of ninety three older adults over 60 years of age (82 women and 11 men; mean age 80.7, SDS 8.91) were recruited from two different backgrounds.

The first sub-sample included fifteen active older adults (9 women and 2 men; mean age 73.7; SD 5.95) who lived independently in the community and volunteered to participate in the study. At the time of data collection, all participants were active members of either Sokol Dolni Chabry or GEMA, both of which are organizations providing physical, cultural, or social activities for older adults. This sub-sample was considered as higher functioning and was also used for the pilot study purposes.

The second sub-sample included seventy eight participants (69 women and 9 men; mean age 81.9; SD 9.31) who were recruited from a larger study 'Dance and Quality of Life' (Vankova, Holmerova, Andel, Veleta, & Janeckova, 2008) which was conducted in Residential Care Facilities (RCF) between 2005-2007. From the total of seven RCFs, additional data were collected in three RCFs between October 2006 and June 2007. RCFs are a common type of long-term care setting in the Czech Republic as well as in other European countries. Although functional limitations are not a condition for admission in the Czech Republic (I Holmerova, 2007), residents in RCFs tend to be more functionally impaired than the rest of the Czech older population (Kalvach, Janeckova, & Bures, 2004). This sub-sample was considered as lower functioning.

All participants from the second sub-sample signed an informed consent form. The Ethical Committee and the Internal Review Board at the Internal Grant Agency of the Ministry of Health of the Czech Republic approved the study. From the total of ninety eligible participants, seventy agreed to participate in the study and another eight participants volunteered to participate even though they were not recruited for the dance study. Ninety-three participants met the eligibility requirements: over 60 years; oral agreement with performance testing; health conditions that would allow physical burden; and basic mobility ability assessed by expert evaluation during an interview prior to the performance testing. Individuals, who were advised by their doctors not to exercise because of existing medical conditions, did not participate in the study.

3.5.2 Procedure

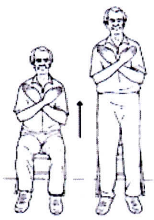
All testing was done in the field setting such as participant's room, activity room, hall, or exercise room. Prior to performing the activities in the SFT battery, all procedures were explained by the administrator and series of warm up activities were performed by the participants. Warm-up activities included marching in place while seated, swinging the arms and stretching large-muscle groups. Participants were asked to do the best they could on all tests but never to push themselves to a point of overexertion or beyond what they thought was safe. The following section will offer a full description of each test within the SFT battery including the equipment requirements, test procedures and scoring instructions (Rikli & Jones, 2001), and the adaptations made in this study.

During testing, participants were closely monitored for signs of overexertion and were discontinued from the study if they were experiencing any of the following symptoms: unusual fatigue or shortness of breath; dizziness; tightness or pain in the chest; irregular heartbeats; pain of any kind; numbness; loss of muscle control or balance; blurred vision; or confusion and disorientation.

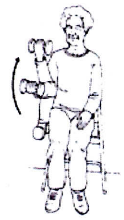
3.5.3 Clinical instruments

Senior Fitness Test (Rikli & Jones, 2001)

The SFT was designed to be easily administered without extensive time, equipment, or space requirements. The test battery can be given by a trained instructor in approximately 15 to 20 minutes to an individual or in approximately 30 minutes to a small group (up to five participants). All of the tests were standardized for older population over 60 years of age and satisfactorily reliable. A detail description of each test used in the study is below:


30-second chair stand test

- Equipment: stopwatch and standard chair (height of 43cm) placed against the wall to prevent slipping
- Procedure: the participant was instructed to sit in the middle of the chair with the back straight, feet flat on the floor, and arms crossed at the wrist and held against the chest. On the signal 'start' the participant rose to a full stand, then returned to a fully seated position. The participant was encouraged to complete as many full stands as possible in 30 seconds. The test was demonstrated slowly to illustrate the proper form, than at a faster pace to show that the object is to do the best one can within safety limits. Before testing the participant practiced one or two stands to ensure the proper form.
- Scoring: the score was a total number of stands completed in 30 seconds. Only one test trial was administered.
- Adaptations: if participant could not perform even one stand without using their hands, he or she was allowed to push off by his or her legs or by chair arms (if any). The adaptation was described but the score for the purposes of this study was zero.



30-second arm curl test

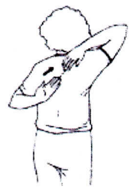
- Equipment: stopwatch, standard chair (height of 43cm) and dumbbell (2.27kg for women and 3.63 kg for men)
- Procedure: the participant was in seated position with the back straight and feet flat on the floor, and with the dominant hand close to the edge of the seat. The weight was held down at the side in the dominant hand with a handshake grip. From the down position, the weight was curled up with the palm gradually rotating to a facing-up position during the flexion. The weight was then returned to the fully extended position with the handshake grip. On the signal 'start' the participant curled the weight through the full range of motion as many time as possible in 30 seconds. The upper arm had to remain still during the test. Bracing the elbow against the body helped stabilize the upper arm. The test was demonstrated slowly to illustrate the proper form, then in a faster pace to illustrate the pace. The participant practiced one or two repetitions without and with the weight to ensure the proper form.
- Scoring: the score was a total number of arm curls completed in the 30 seconds. If the arm was more than halfway up at the end of 30 seconds, it counted as a curl. Only one trial was administered.
- Adaptations: slightly lighter dumbbells were used (2.5 kg for women and 3.5 kg for men). The reason was that the original dumbbells were not available in the Czech Republic.

2-minute step test

- Equipment: stopwatch, piece of string and masking tape
- Pre-procedure: the minimum stepping height is at the level even with the midway point between the kneecap and the front hip bone. The proper height of steps was determined using a tape measure. The correct height of knee was corrected by moving the participant to the wall or a door way and transferring the tape to a spot at the same level on the wall or between the door.
- Procedure: participant was asked to stand against the wall or door way and on the signal 'start' to begun stepping in a place as many times as possible in the 2-minute period. Although both knees had to be raised to the correct height, only the number of times the right knee reached target was counted. When the proper knee height could not be maintained, the participant was asked to slow down, or to stop until he or she could regain the proper form, but the time was kept running. The test was demonstrated slowly to illustrate the form, then in a faster pace to illustrate the pace. The participant practiced four to six repetitions to ensure proper form.
- Scoring: the score was the number of full steps completed in 2 minutes. Only one trial was administered.
- Adaptations: if participant was unstable, he or she was allowed to hold the wall or chair. If participant was unable to lift their knees to a proper height or could lift only one, he or she was allowed to complete the test, but it was indicated on the score card and zero score was administered for the purposes of final analysis.

 Chair sit-and-reach test

- Equipment: standard chair and ruler
- Procedure: participant was seated on the edge of the chair. He or she extended one leg as straight as possible in front of the hip with the heel on the floor and foot flexed at approximately 90 degrees. The other leg was bent with the foot flat on the floor. With arms outstretch, hands overlapping, and middle fingers even, the participant slowly bended forward at the hip joint reaching as far as possible towards or past the toes. If the extended knee started to bend, the participant was asked to move slowly back until it was straight. The maximum reach must have been held for two seconds. Only the preferred leg was used for scoring purposes. The test was demonstrated and participant practiced one or two times to ensure proper form.
- Scoring: the distance from the tip of the middle fingers to the top of the shoe was measured to the nearest half centimeter. Minus points were given if the middle fingers did not touch the top of the shoe, zero score if the middle fingers just touched, and plus scores if the middle fingers overlapped. Better of two scores was administered.



Back scratch test

- Equipment: ruler
- Procedure: participant was asked to stand and place his or her preferred hand over the same shoulder, palm down and fingers extended, reaching down the middle of the back as far as possible. Then the participant was instructed to place other hand around the back of the waist with palm up, reaching up of the middle of the back as far as possible in an attempt to touch or overlap the extended middle fingers of both hands. The test was demonstrated and participant practiced one or two times to ensure proper form.
- Scoring: the distance between the tips of both middle fingers was measured. Minus points were given if the middle fingers did not touch, zero score if the middle fingers just touched, and plus scores if the fingers overlapped. Better of two scores was administered.

8-foot up-and-go test

- Equipment: stopwatch, standard chair, tape measure and cone
- Setup: the chair was placed by the wall if possible facing a cone marker exactly 2.44 meters (8 feet) away.
- Procedure: participant was instructed to sit in the middle of the chair with back straight, feet flat on the floor, and hands on the thighs. One foot should have

been slightly in front of the other foot, with the torso slightly leaning forward. On the signal 'start' the participant got up from the chair, walked as quickly as possible around either side of the cone, and then sit back down in the chair. The test was demonstrated and participant practiced one time to ensure the proper form.

- Scoring: the trial was administered to the nearest second.
- Adaptations: if needed, a cane or a walker could have been used for this test but the score was not considered for the final analysis.

3.5.4 Data analysis

The purpose of this study was to verify the applicability of SFT among older Czech adults living in the Czech Republic either independently or in RCFs and describe in detail functional fitness levels of both sub-samples. Additionally, actual functional fitness status was compared between and within both sub-samples and the total percentage of older adults who performed at or below the risk of independency loss was detected. The SPSS 16 program for Widows (SPSS Inc., Chicago, IL, USA) was used to examine all variables. Means and standard deviations were performed to describe both sub-samples. Independent-samples *t*-test and one-way ANOVA were used to determine the differences between and within the sub-samples.

3.6 Results

Applicability of Senior Fitness Test in the Czech Republic

All participants from home settings completed all tests included in the SFT battery without any problems. But participants living in Residential Care Facilities experienced some difficulties with completing SFT battery. The summary is presented in the table 3.1. It appeared that the most difficult test was the 8-foot up-and-go test (22% participants found it too difficult to perform) followed by the 2-minute step test (21% participants found it too difficult to perform). Conversely, the easiest test was the 30-

second arm curl test (only 1 participant was not able perform). Both of the flexibility tests: the chair sit-and-reach test and the back scratch test, and the 30-second chair stand test happened to have medium difficulty (12%, 8%, and 6% participants found it too difficult to perform). Women had more problems compared to men, but it could be caused by unequal sample size. When we divided women into three different age categories as recommended by Holmerova & Juraskova (2003), it appeared that with increasing age participants experienced more difficulty. 80% women in the youngest category were able to perform complete SFT compared to just 55% women from the middle category. The oldest category (over 90 years) experienced even more difficulties when compared to the middle one. Only 50% of the women from that category were able to complete the test battery.

Table 3.1 Total number (percentage) of residents of Residential Care Facilities who were not able to complete individual Senior Fitness Test tests for the total sub-sample, for women and men separately, and for women only * divided into three age categories

	All (N =78) N (%)	Women (N = 69) N (%)	Men (n = 9) N (%)	60-74* (n = 10) N (%)	75-89* (n = 49) N (%)	90 plus* (n = 10) N (%)
Test 1: 30-second chair stand test [number of repetitions]	5 (6%)	5 (7%)	9 (0%)	0 (0%)	2 (4%)	3 (30%)
Test 2: 30-second arm curl test [number of repetitions]	1 (1%)	1 (1%)	9 (0%)	0 (0%)	1 (2%)	0 (0%)
Test 3: 2-minute step test [number of repetitions]	16 (21%)	14 (20%)	2 (22%)	0 (0%)	9 (18%)	5 (50%)
Test 4: chair sit-and-reach test [cm]	9 (12%)	9 (13%)	9 (0%)	0 (0)%	6 (12%)	3 (30%)
Test 5: back scratch test [cm]	6 (8%)	6 (9%)	9 (0%)	0 (0%)	5 (10%)	1 (10%)
Test 6: 8-foot up-and-go test [seconds]	17 (22%)	16 (23%)	1 (11%)	2 (20%)	9 (18%)	5 (50%)
Complete SFT (non-valid N listwise)	**31 (40%)**	**29 (42%)**	**2 (22%)**	**2 (20%)**	**22 (45%)**	**5 (50%)**

Differences between older adults living in Residential Care Facilities and those living in home settings

Differences in performances between those living in Residential Care Facilities and those living in home settings were observed for each test within the SFT battery. These differences were clinically significant for all tests (figure 3.1) and statistically significant for five out of six tests (table 3.2). Independent *t*-tests were performed to examine whether means for the both sub-samples were significantly different from zero. Except for the chair sit-and-reach test (t (82) = -1.64, p = .104), all of the differences in mean performance were highly statistically significant: 30-second stand test (t (86) = -5.05, p = .000); 30-second arm curl test (t (90) = -5.24, p = .000); and 2-minutes step test (t (75) = -6.57, p = .000). The differences in the last two tests were also statistically significant, but because of observed inequality of variances, a different *t*-test statistic was performed. The Levene's test of homogeneity of variance was used to verify equal variance assumption. The null hypothesis for the Levene's test is that the variances are homogeneous. For this set of data the Levene's test was not significant for most of the tests, indicating the null hypothesis cannot be rejected, that is, the variances are homogeneous, except for the back scratch test (F = -2.009, p = 0,011) and the 8-foot up-and-go test (F = 6.405, p = .014). Thus the Welsh's *t*-test statistic for non equal variances was used instead and the differences in mean performance were also highly statistically significant: back scratch test (t (66) = -3.56, p = .000) and 8-foot up-and-go test (t (62) = 4.95, p = .001).

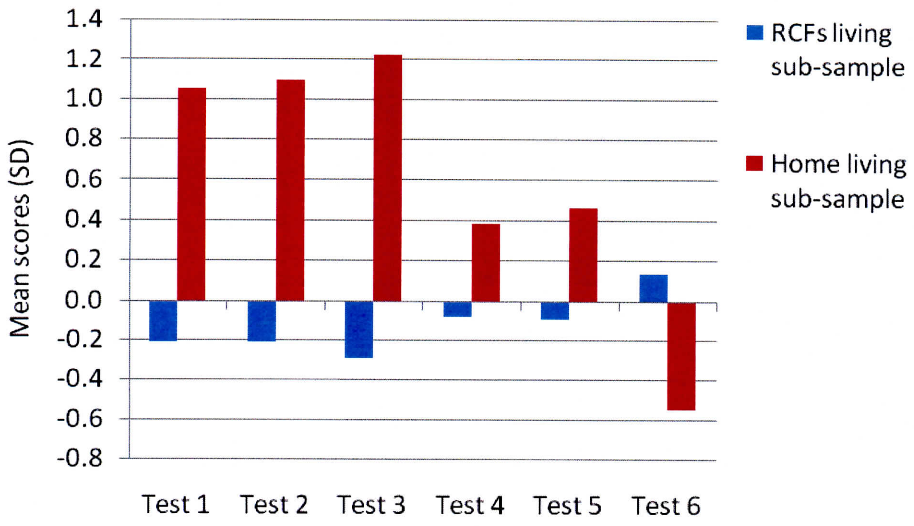


Figure 3.1 Mean performance in Z-scores for each Senior Fitness Test test for all participants (both women and men) living in Residential Care Facilities settings (blue) and home settings (red)

Table 3.2 Mean performance in original units and standard deviations for each Senior Fitness Test test for both sub-samples and *t*-test statistic with statistical significance indicating the differences between means

	Home living sub-sample (n = 15) mean ± SD	RCF sub-sample (n = 78) mean ± SD	*t* (df)	Sig.
Test 1: 30-second chair stand test [number of repetitions]	13.5 ± 4.0	8.2 ± 3.7	-5.05 (86)	.000
Test 2: 30-second arm curl test [number of repetitions]	19.9 ± 5.0	12.6 ± 4.9	-5.24 (90)	.000
Test 3: 2-minute step test [number of repetitions]	76.7 ± 24.9	34.9 ± 21.4	-6.57 (75)	.000
Test 4: chair sit-and-reach test [cm]	-0.4 ± 7.6	-4.9 ± 10.0	-1.64 (82)	.104
Test 5: back scratch test [cm]	-11.7 ± 6.3	-21.2 ± 18.0	-3.56 (66)*	.000
Test 6: 8-foot up-and-go test [seconds]	5.9 ± 1.5	20.8 ± 23.4	4.94 (62)*	.001

Note: * Welch's *t*-test statistic for not equal variances

The role of age on the performance of older women living in Residential Care Facilities

Three age categories for older adults recommended by Holmerova & Juraskova (2003) were used: the youngest category (60-74 years – category 1); the middle category (75-89 years – category 2); and the oldest category (90 years and over – category 3). The number of women in each category varied from ten women in the youngest and the oldest category to forty nine women in the middle one. As seen in the figure 3.2, except for the chair sit-and-reach test, we can certify that physical fitness declines with age and that this decline is visually obvious. To examine the statistical significance of this decline, we performed one-way ANOVA. This approach is recommended when there is a need to test differences between more than two groups. It appeared that only the difference in the 30-second chair stand (F (63) = 4.85; p = .011) was statistically significant. Otherwise, the observed decline was not statistically significant but may be considered as clinically significant. The F-test statistics was: for the 30-second arm curl test F (67) = 0.95, p = .393; for the 2-minute step test F (54) = 1.43, p = .249; for the chair sit-and-reach test F (59) = 1.00, p = .374; for the back scratch test F (62) = 1.54, p = .222; and finally for the 8-foot up-and-go test F (52) = 0.91, p = .408 (table 3.3).

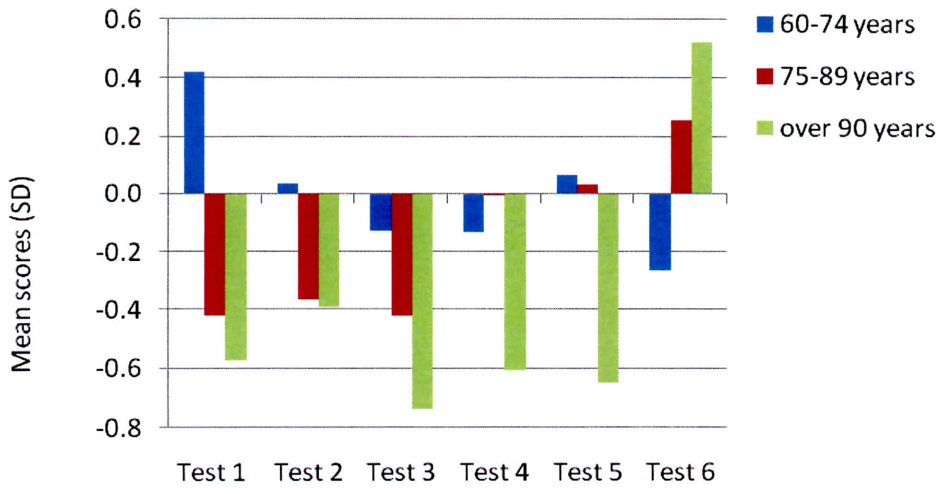


Figure 3.2 Mean performance in Z-scores for each Senior Fitness Tests test for women living in Residential Care Facilities divided into three different age categories (60-74 years in blue, 75-89 years in red, 90 years and over in green)

Table 3.3 Mean performance in original units and standard deviations for each Senior Fitness Test test for women in all three age categories, and *F*-test statistic with statistical significance indicating the differences between means

	60-74 (n = 10) mean ± SD	75-89 (n = 49) mean ± SD	90 plus (n = 10) mean ± SD	*F*(df)	Sig.
Test 1: 30-second chair stand test [number of repetitions]	10.9 ± 4.3	7.3 ± 3.3	6.7 ± 2.6	4.85(63)	.011
Test 2: 30-second arm curl test [number of repetitions]	14.0 ± 6.1	11.8 ± 4.7	11.6 ± 4.1	0.95(67)	.393
Test 3: 2-minute step test [number of repetitions]	39.5 ± 21.3	31.5 ± 17.7	22.8 ± 21.1	1.43(54)	.249
Test 4: chair sit-and-reach test [cm]	-5.4 ± 10.8	-4.1 ± 9.3	-10.0 ± 14.9	1.00(59)	.374
Test 5: back scratch test [cm]	-18.4 ± 18.3	-19.0 ± 17.9	-30.6 ± 21.6	1.54(62)	.222
Test 6: 8-foot up-and-go test [seconds]	12.2 ± 3.8	23.5 ± 27.1	29.3 ± 22.3	0.91(52)	.408

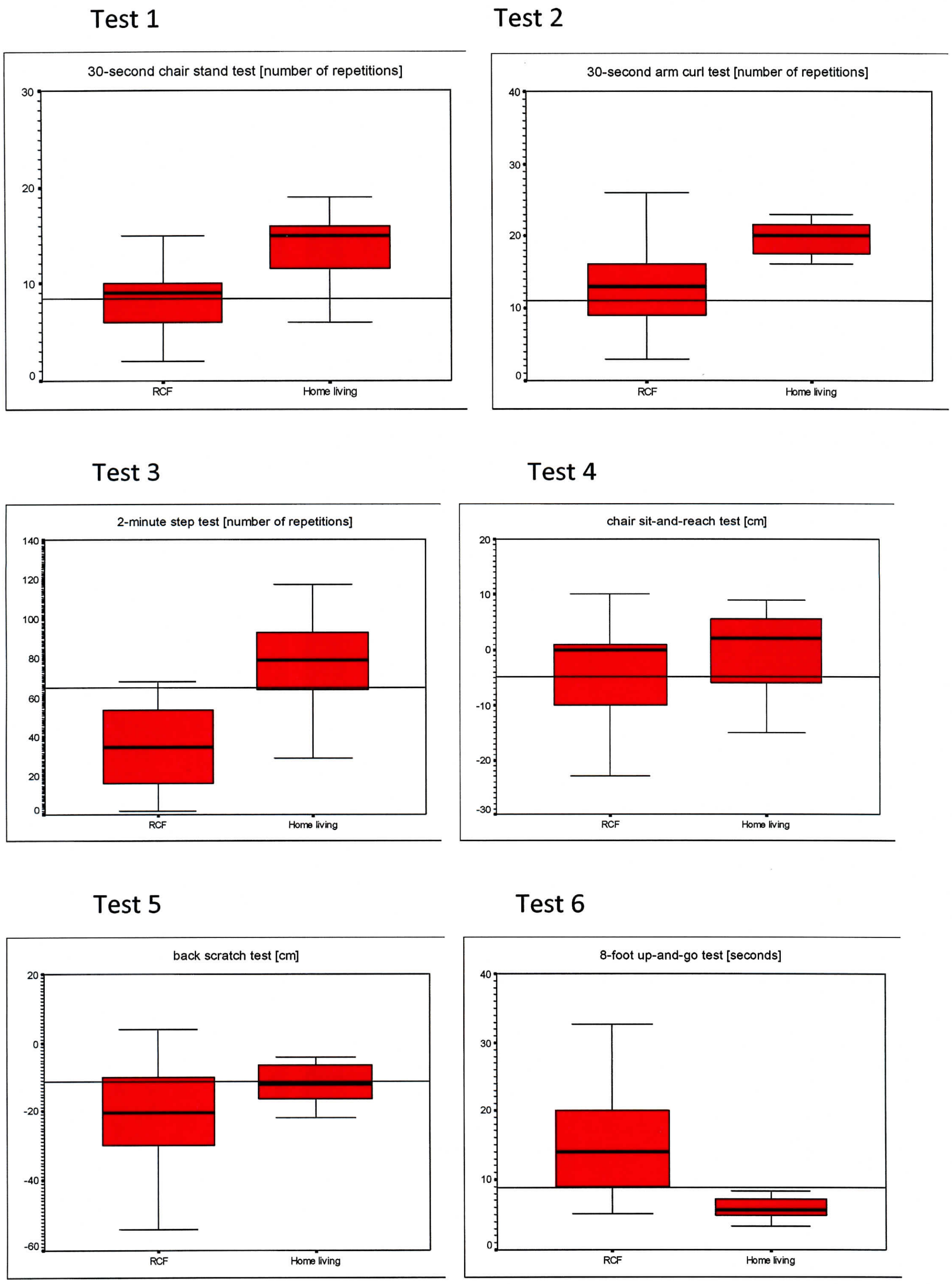
Physical fitness performance and the risk of independency loss

Physical activity has been recognized as an effective prevention for maintaining independence. The goal of many physical therapists (among others) is to be able to successfully detect patients who are at risk of becoming dependent in order to effectively preserve functional independence through appropriate intervention programs. To detect those older adults who are performing dangerously close to the minimal level needed for participating independently in daily activities, threshold fitness scores were developed by Rikli & Jones (2001). The additional goal of the authors of the SFT was to identify scores associated with the risk for losing one's ability to function independently. So in addition to testing participant's fitness levels, they also assessed the functioning ability levels. Specifically, functional ability was assessed through self-evaluation using a composite physical functional scale that had been previously standardized. The 12-item battery was designed to assess function across wide range of abilities – from basic activities of daily living through intermediate or instrumental activities of daily living to advanced activities such as strenuous household, sports, and exercise activities. The composite score was used to categorize the individuals as either high functioning or low functioning, with high functioning being those who indicated that they can perform all twelve tasks with no difficulty and low functioning being those who reported they can perform no more than six or fewer (50%) of the tasks without difficulty. Of special interest was the fitness level that was associated with being low functioning and potentially at risk for loss of physical independence (Rikli & Jones, 2001). The data in the Rikli & Jones's study revealed a strong positive association between fitness level and self-reported physical ability level. The average fitness score of those who reported having difficulty with common everyday activities associated with independent living provided a type of threshold value that is associated with loss of functional fitness. Table 3.4 presents the average fitness scores associated with low levels of self-reported functional ability for both men and women and the number (percentage) of participants from both sub-samples performing at or below the threshold score (more described below). Figure 3.3 indicates the at-risk scoring zones associated with loss of function for women in both sub-samples graphically. Even though the threshold scores were developed for U.S. populations, we believe that this approach can provide useful, so far in the Czech Republic unavailable, reference points for interpreting physical capacity in Czech older adults and to target the attention on the most needed individuals who are at the risk of independency loss.

Table 3.4 also presents, except for the threshold scores, the total percentage and actual number of participants who performed at or below the threshold score for all functional fitness tests. As indicated in table 3.4 women were more at the risk compared to men. As we expected residents for RCFs were mostly at the risk compared to those who were and maybe still are living at home settings. As in the previous analyses it appeared that the Czech older adults had the most problems with the completion if the 2-minute step test and one of the two flexibility tests - the back scratch test.

Table 3.4 Threshold scores and total percentage (number) of participants of both sub-samples who performed on or under threshold scores associated with being at risk for losing one's ability to function independently

	Threshold scores for each fitness component		Home living sub-sample		RCFs sub-sample	
	Wo-men	Men	Women (N = 13) N (%)	Men (N = 2) N (%)	Women (N = 69) N (%)	Men (N = 9) N (%)
30-second chair stand test [number of repetitions]	8.4	8.3	3 (23%)	0 (0%)	40 (58%)	1 (11%)
30-second arm curl test [number of repetitions]	11.0	10.8	1 (8%)	0 (0%)	33 (48%)	1 (11%)
2-minute step test [number of repetitions]	65	65	5 (38%)	0 (0%)	67 (97%)	7 (78%)
chair sit-and-reach test [cm]	-4.8	-6.3	3 (23%)	1 (50%)	31 (45%)	3 (33%)
back scratch test [cm]	-11.4	-20.3	6 (46%)	0 (0%)	50 (72%)	6 (67%)
8-foot up-and-go test [seconds]	8.8	8.9	0 (0%)	0 (0%)	58 (84%)	7 (78%)



Note: The box-plot is a graphical interpretation of the data using five-number summaries (the smallest number, lower quartile, median, upper quartile, and the highest number)

Figure 3.3 Box-plots for women in both sub-samples illustrating differences between groups with the respect to threshold scores for *each SFT test*

3.7 Discussion

The present study has revealed that the SFT is a very useful measurement instrument to assess functional fitness among older adults at wide range of ability levels. SFT was successfully applied on the Czech older population in order to examine whether this battery is acceptable for the majority of older adults living in the Czech Republic. Performance-based measures consisting of objective observations of functional capacity were claimed to be applicable cross-culturally because they seem less likely to be influenced by culture, language, and educational level when compared to self-report measures (Ferrer, Lamarca, Orfila, & Alonso, 1999). Data about Czech older adults' fitness are sparse and incomplete. This study was conducted to gain a better understanding of functional fitness levels of older adults living in the Czech Republic and desirably, to set up the ground for future research conducted to develop normative tables.

Even though the SFT was developed for community dwelling older adults over 60 years, is has been demonstrated that this measurement tool was applicable on low functioning older adults either living in nursing homes (Beck et al., 2008) or older adults suffering from chronic conditions ((Alexander et al., 2008; Mahrova et al., 2006) as well as on extremely fit older adults (Thompson et al., 2007). In order to confirm that SFT will be accepted by the majority of older adults living in the Czech Republic the study sample was recruited from two quite different backgrounds. The first sub-sample included independently living community residents who were considered as high functioning and on the other hand, the second sub-sample included permanent residents of RCFs who were considered as low functioning. Both of the sub-samples accepted performance testing very well. However, quite a few RCFs' residents, especially those in advanced old age, have some difficulty with performing some of the tests compared to independently living sub-sample where none of the participants experienced any difficulties with none of the test.

It was not surprising that older adults from the youngest category (60 to 74 years) experienced less difficulty that older adults form the middle category (75-89) and that the oldest older adults (over 90 years) found the SFT quite difficult. But most importantly, still the half of the oldest ones was able to perform complete SFT. The SFT was accepted by the majority of screed population, therefore we can consider this battery generally applicable even for low functioning older adults. **The hypothesis H1 was accepted**. The fact that some of RCFs' residents scored 0 in some of the tests reflects truly their actual fitness status and provides the valuable information. For instance, after an intervention program with low functioning older adults, the improvement from 0 to 5 may be detected. This improvement will be significant clinically and most likely even statistically.

Although the functional limitation is not a condition for admission in the Czech Republic (I Holmerova, 2007), residents in RCFs tend to be more functionally impaired than the rest of the Czech older population (Kalvach et al., 2004). Also, residents in RCFs do not need to take care of themselves in the sense of being responsible for instrumental activities of daily living (grocery shopping, running errands, cooking, housekeeping and many others), so their activity is very often limited to just basic activities of daily living which may cause an even faster decrease in physical functioning resulting in actual lower levels of functional fitness. On the other hand, older adults living still independently in their homes are forced to accomplish IADL, therefore they were considered as higher functioning. Additionally, all of them were active members or leisure activity organizations (GEMA or Sokol). Based on these facts, we hypothesized, that the sub-sample recruited from RCFs will have significantly lower functional fitness compare to more active older adults who are still living independently and participating in leisure time activities. Statistical differences between those two sub-samples were significant for all of the tests except for the test 4: chair sit-and-reach test where the difference was just clinical. According to our results **the hypothesis H2 was accepted**.

All components of functional fitness decrease with age in all fitness components (Frankel et al., 2006; Rikli & Jones, 2001; Spirduso, 2005). However, which components are more sensitive to age effects is not clearly understood, especially in the Czech Republic. Because of the number of participants living independently was low and the gender distribution was unequal, we were able to analyze only women living in RCFs. Age-related decline was statistically significant only for the test 1: 30-second chair stand test. All mean scores rapidly decreased between the youngest and the oldest category. The smallest relative decline was observed in the test 3 measuring upper body strength (17%). Conversely the biggest relative decline was observed in the test 6 measuring dynamic balance/agility (140%). The decline in the rest of the tests varied between 38% to 85%. The hypothesis **H3a was accepted** because the decline was considered clinically significant for all tests.

Another question was if the decline remains linear in old age (between older individuals from their 60s to their 90s). In previously published studies, the focus was targeted on the lifespan changes rather than changes when one is already in old age and most attention was paid to aerobic endurance or muscular strength. As demonstrated by Wilson and Tanaka (2008) in their meta-analysis the decline for aerobic capacity is linear through the lifespan (Wilson & Tanaka, 2000). This finding was supported in our study. The relative decline in aerobic capacity was the same across age groups: 20% decline was observed between the youngest category and the middle category and 22% decline was observed between the middle group and the oldest group. We expected the same scenario for the rest of the tests but it appeared that the decline slowed down with advanced old age. The most remarkable changes were observed between the youngest and the middle category in three out of six tests (except for the aerobic endurance and both flexibility tests). The decline varied from 15% to 92% while additional decline between the middle category and the oldest category slowed down rapidly (5% between the middle group and the oldest group versus 33% the youngest group and the middle group; 1.5% versus 15%, and 48% versus 92%). This trend did not appear in the aerobic endurance test where the decline continued by the same rate and in both flexibility tests. The hypothesis **H3b was accepted only for test 3: the 2-minute step test and rejected for the rest of the tests**. According to our results it seems that the decline in muscle strength and agility slows down with the increasing age. In other words, the most dramatic changes were observed between the age 60-74 and age 75-89. There is a need for more studies which would support these findings.

The ability to recognize the "at risk of independency loss" individuals ahead of time is essential. But what is the risk level? So called threshold scores indicating the risk levels have been developed by Rikli and Jones (2001). Since there are no available such scores for use in the Czech Republic, the U.S. version was transferred. The rationale for such an approach was solely logical: the physical capacity measured by the performance test as functional fitness should be the same for anybody in order to be independent in activities of daily living despite the cultural differences. For instance the same lower-body strength would be needed for anybody with the same proportions (height, weight and length of extremities) to rise from the chair without any additional help. The level where one will not able to rise will be the same as well. According to the theory, it was expected that the independent sub-sample will score above the threshold level in all tests by all participants and that more than 50% of RCFs residents will score below. **The hypothesis H4a was rejected and the hypothesis H4b was accepted**. This means that the threshold scores for the Czech population might be less strict especially for the test 3: 2-minute step test and the test 5: back scratch test for women. The reason might be that even logically the threshold scores were expected to be the same across cultures they are not because they were developed using self-reports measures (for details see the 'results' section above) which may be influenced by population biases. Basically, the threshold scores were mean scores for the sample based on the subjective rating of the functional ability which might differ across countries and cultures. Until there are proper threshold scores developed for the Czech Republic, the U.S. version may be used with caution and solely for the overall idea and participants' motivations.

Several limitations should be mentioned. First, the sample consisted of mostly women so we were not able to explore whether some of our findings apply to both genders. Second, the mean age of both sub-samples was not the same but it did not influence the significance of the results. The analysis was replicated with an incomplete RCFs' sub-sample (participants over 85 years of age were excluded) so the age range was exactly the same and it appeared that the *t*-test statistic was still highly significant. Third, the total number of participants was quite low and also the number of participants was not alike in each age category thus additional research is needed to support our findings. Fourth, the all the participants volunteered so conclusions cannot be generalized. And finally fifth, the threshold scores might have been biased by the cultural differences so they should be used with caution. Also the dumbbells used in this study were not exactly the same because of different metric system. The differences were considered minor (2.7% lighter for men and 8.7% lighter for women) which most likely could not significantly influence the results. Nevertheless it would be desirable to use exactly the same dumbbells for future research.

In conclusion, the SFT was well accepted by all tested older adults living in the Czech Republic. Even though a few participants experienced difficulty with some of the tests, it did not cause either floor or ceiling effects. Therefore this test battery was considered to be a great instrument measuring functional fitness and it is recommended for future research conducted to develop normative tables for older Czech adults. As expected, independently living older adults have better functional fitness compare to those living in RCFs so the battery seems to be sensitive enough to detect differences in fitness status. Also, all functional components declined with age as was illustrated in women RCFs' sub-sample which supports the sensitivity as well. Additionally, it appeared that the age-related decline in the same sub-sample slowed down with increased age except for aerobic endurance where the decline remained the same between all three groups and for flexibility tests where the decline did not follow this trend. Although the present dissertation is generally about the importance of functional fitness, so far we were able to analyze just its individual components and describe them only individually. It would be useful and interesting to be able to measure overall functional fitness besides analyzing just its component separately. The composite measurement score would be valuable for both clinical and research purposes because the multiple evaluations are not always satisfactory or even desirable. But because the functional fitness is a latent - in other words unobservable construct, development of composite score requires a special attention and procedures. The question what would be the most accurate estimation of the overall functional fitness will be addressed in the next chapter.

CHAPTER 4 – Overall Functional Fitness model for older adults

4.1 Abstract

Background: Functional fitness is a latent construct consisting of specific components essential for independent functioning; including muscle strength, stamina, flexibility and agility. Accurately estimating overall fitness is essential for clinical and research purposes because multiple evaluations are not always satisfactory or even desirable. However, the structure of functional fitness and the contribution of its individual components have been under-investigated.

Methods: A 6-test performance battery (Senior Fitness Test - SFT) was administered to seventy eight participants (69 women and 9 men; mean age 82.0; SD 8.8). A single level structure model of Overall Functional Fitness was tested. Six indicators represented by 30-second chair stand test, 30-second arm curl test, 2-minute step test, chair sit-and-reach rest, back scratch test, and 8-foot up-and-go test were hypothesized to have loadings on the first-order factor represented by the Overall Functional Fitness. The structural equation modeling using Lisrel statistical package was performed to test the hypothesis.

Results: Standardized maximum likelihood solution of the initial model showed satisfying goodness-of-fit indices (RMSEA = 0.00; NFI = 0.95; GFI = 0.97; RMR = 0.038 and $\chi^2 = 7.64$, df = 9). The factor loadings of the manifest variables on the first-order factor were greater than 0.54. Findings demonstrate that the structure of functional fitness is unidimensional and contributions of its individual components are hierarchical in nature.

Conclusion: An accurate estimation of Overall Functional Fitness considers the weighted sum, rather than a simple sum, of all identified components. Using an estimation of Overall Functional Fitness contributes to implementing and evaluating effective activity/exercise programs targeting the select needs of older individuals that become increasingly variable with the onset of diseases or disabilities that often accompany the aging process.

4.2 Introduction

Rationale for evaluation of composite measurement score in general

Physicians and medical researchers are faced with the necessity of measuring complex phenomena or constructs such as disease risk or severity, physical disability, functional fitness, or quality of life. The development and psychometric evaluation of scales which measure unobservable - latent constructs continues to be an issue of high interest among many researchers. Meaningful summary measures are needed because the use of multiple evaluations is not always satisfactory or even desirable.

Although some of these complex approaches may be considered as global holistic opinions, another approach based on the use of the composite measurement score (CMS) may be adopted. CMS are increasingly used in medicine to measure complex constructs in the absence of a reference criterion or 'gold standard'. Briefly, the construct is considered to be composed of several homogenous dimensions which are evaluated by a number of selected elementary tests. Each such construct is typically represented by multiple manifest variables that serve as indicators. This is called a structure equation model which is in other words hypothesized pattern of directional and non-directional linear relationships among a set of manifest variables and latent variables. To characterize subjects, detailed scores of manifest variables are combined into one composite measurement score.

However, a review of current practice in the construction of evaluative composite measurement scores showed serious flaws that mainly result from a lack of basic measurement properties and requirements. A weighting system proposed to 'adjust' the relative contribution of indicators should be also considered when creating CMS.

Development of composite measurement scores

The development of a composite measurement score of a latent construct requires validation of the hypothesized relationship between the construct and its indicators. The indicators are typically single items or composites consisting of multiple items. The whole process involves complex procedures described in detail below and requires an appropriate methodology. Important measurement properties such as measurement level, content and construct validity, and reliability are needed to be remained (Coste, Fermanian, & Venot, 1995; Coste, Walter, & Venot, 1995). The composite measurement scores are not informative but practical and important for many research and clinical purposes not only in gerontology.

Measurement consists of the rules for assigning the numbers to objects in such a way as to represent quantities of attributes. An attribute is a particular characteristic of object according to which they can be ordered. For easily definable attributes such as height, weight or urea concentration, the rules are obvious and detailed formulations are not necessary. By contrast, the rules measuring constructs, that are hypothesis constructed by scientist to explain relationship between observables (behavior, depression, disability, fitness etc.), are not so obvious. A composite measurement score is the result of the process of construction from simple raw values for tests to a single score.

Scales and levels of measurement: scales are the results of the measurement procedures and can be classified according to levels of measurement – nominal scale, ordinal scale, interval scale and ratio.

Basic requirements of measures - psychometric properties: according to psychometric theory, the basic requirements of measures are: validity, reliability and sensitivity to a change. An essential, but often overlooked, property of measurement which is assumed in both exploratory and confirmatory statistical technique is unidimensionality. Scales which are unidimensional measure a single trait. This property is absolutely essential for un-confounded assessment of variable interrelationship in the modeling. The composite measurement score is meaningful only if each of the items is acceptably unidimensional.

The mathematical definition of unidimensionality is based on traditional common factor model in which a set of indicators share only a single underlying factor. Development and evaluation of measurement scales has traditionally relied upon analyses such as coefficient alpha, item-to-total correlations, and exploratory factor analysis. The use of confirmatory factor models for such purposes is relatively new phenomenon and has its history of application within the areas of education and psychology (Bentler, 1986). Importantly, only confirmatory factor analysis (CFA) directly tests unidimensionality as formally defined in the equations developed earlier. In the other words, the CFA provides direct and quantifiable evidence regarding the external and internal consistency among a set of construct indicators. A formal definition of the CFA model is contained in Joreskog and Sorbom's LISREL model (Joreskog & Sorbom, 1989).

The overall fit of the hypothesized model can be tested by using the maximum likelihood χ^2 statistic provided by the LISREL output. Importantly, this measure of fit is a function of external and internal consistency (the difference between observed correlations and those implied by model's estimated parameters). In a general sense, higher χ^2 values are indicative of better fitting models. The χ^2 statistics is sensitive with respect to either small or large sample sizes (Bollen & Long, 1993; Joreskog, 1969) and models with large numbers of indicators (Bearden, Sharma, & Teel, 1982). In these instances even trivial discrepancies between a model and data can result in significant χ^2 values. Therefore, other measures of model fit such as root mean square error of approximation (RMSEA), goodness of fit indices (GFI), normed fit index (NFI), or root mean square residuals are recommended to consider (RMR) (Bollen & Long, 1993; Segars, 1997).

Existence of either of these conditions may confound the results of subsequent modeling and therefore the validity of reported findings. Importantly, theoretically derived models must themselves be grounded in theory. In the absence of theoretical grounding or data recollection, fit of the model may be due to statistical chance rather than empirical representation of underlying phenomena. Any composite measure development should be always done by two (theoretical and statistical modeling), or if necessary, by three steps (grounded re-specification):

1. Theoretical measurement modeling

Many variables of interest are inherently complex in nature and they cannot be accurately measured with a single scale. Additionally, single items tend to frame a concept narrowly resulting in considerable high measurement error. As shown in figure 4.1, the initial step in developing these multi-items measures involves specifying the domain of the constructs. Such a definition is structured through intense review of the relevant literature. The next step in theoretical measurement modeling is the generation of a sample of items for each construct of interest. This should be accomplished through analysis of existing measurement scales, relevant literature, or expert opinion. It is recommended to adopt an existing measurement scale for research purposes. New measurement scales make it difficult to compare and accumulate findings, thereby inhibiting synthesis of what is known. Furthermore, this step insures the completeness of construct operationalization. Strong theoretical underpinnings are required for proper model development and further testing (Segars, 1997).

2. Statistical measurement modeling

The success in capturing latent phenomena though operationalized test scales is assessed through statistical modeling. The initial step of this process is the estimation of a confirmatory factor model. A statistical package such as LISREL may be utilized for this purpose. Upon estimation of the measurement model, a formal assessment of convergent validity can be performed. As noted earlier, examination of fit indices and indicators' loadings provide the researcher with specific evidence regarding these important measurement characteristics (Segars, 1997).

3. Grounded re-specifications

In many instances, fit indices will suggest that improvements in model fit may be gained by eliminating tests with low reliabilities, re-specifying to load on more than one factor, adding additional factors, or correlating error terms. Improvement in measurement may be due to statistical chance rather than substantive gain in empirical understanding of the underlying latent phenomena. As implied in the figure 4.1, it is critical that re-specifications are grounded in and guided by substantive theoretical reasoning. Re-specification of measurement models should not turn into 'treasure hunt' in which numerous configurations are tested until good fitting model emerges (Segars, 1997).

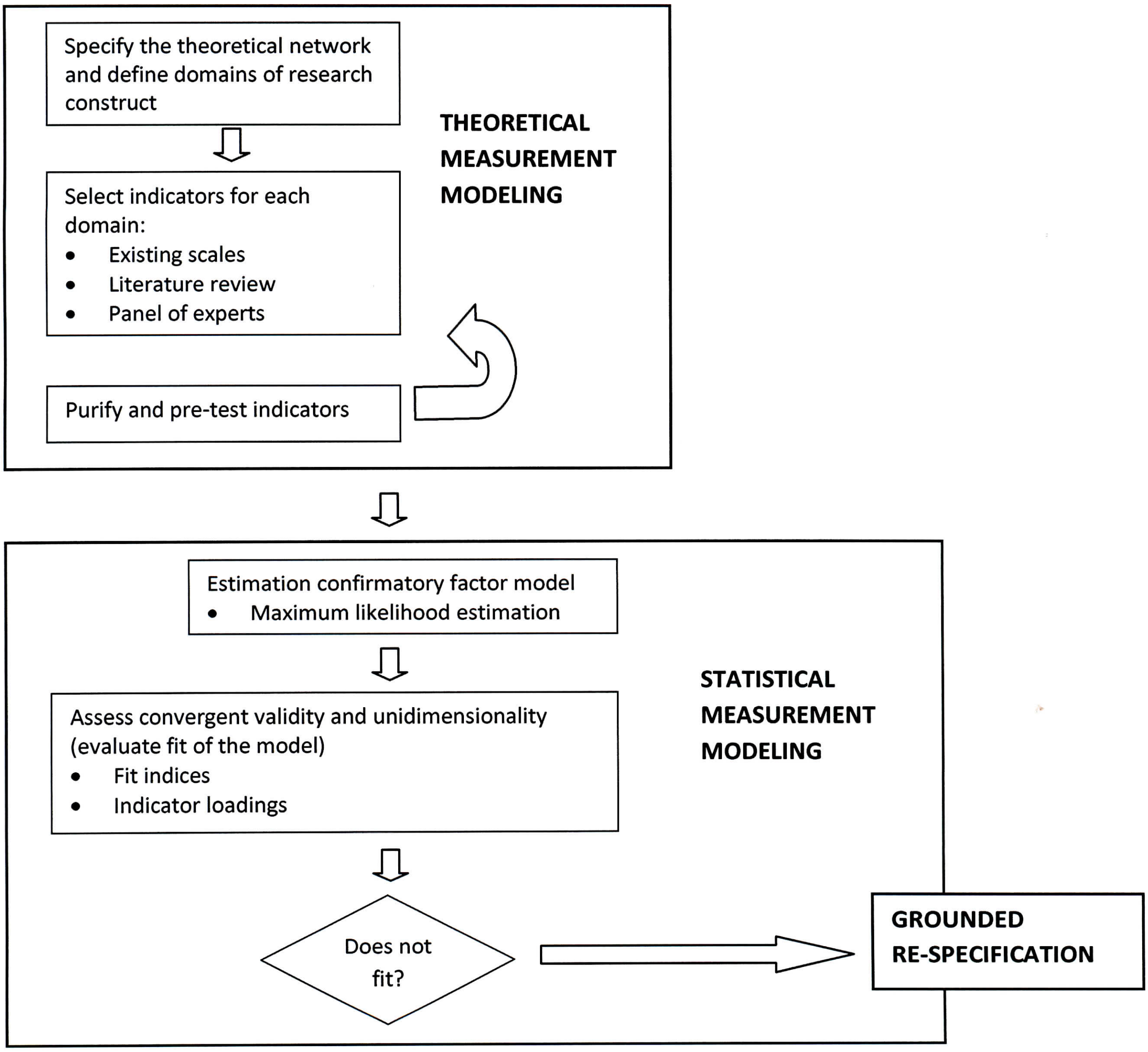


Figure 4.1 Schema of theoretical and statistical paradigm for unidimensional measurement

Evaluations of goodness-of-fit indices

The empirical assessment of the proposed models is a vital aspect of the theory development process, and central to this assessment are the values of fit indices obtained from the analysis of a specified model. A principal source of evaluation of the fit indices is Monte Carlo research (Gerbing & Anderson, 1993), which investigates the distributional properties of statistics by repeated sampling from simulated distributions, usually with the underlying population model known a priori. However, the inherent limitations of Monte Carlo research necessitate a consideration of the definitional and conceptual properties of the indices as well. Even though, the interpretability of the model can be judged only subjectively and it is not amendable to the application of statistical method, it does not render this characteristic of model any less important: it is only more difficult to investigate. The applications of the analysis of covariance structures in the behavioral sciences are implausible, resulting in the fact that any used model is anything more than an approximation to reality. Statistical goodness-of-fit tests are often more a reflection on the size of the sample than on the adequacy of the model (Bollen & Long, 1993).

Type of fit indices

Chi-square test ($\chi 2$) is one of the most widely used theoretical probability distributions in inferential statistics, e.g., in statistical significance tests. It tests for goodness of fit of an observed distribution to a theoretical one. The chi-square statistic should not be the sole basis for determining a model fit because: there is no allowance made for the approximate nature of virtually all behavioral science models; it ignores statistical power of the test; and failure of the variable to satisfy the distributional assumptions of the test statistics can lead to a rejection of correct models (Bollen & Long, 1993).

RMSEA (Root Mean Square Error of Approximation) takes into account the error of approximation in the population and asks "How well would the model, with unknown but optimally chosen parameter values, fit the population covariance matrix if it were available?" This discrepancy is expressed per degree of freedom. The evaluation of this index is based on subjective judgment. Generally, a RMSEA of less than 0.05 indicates a good fit, values as high as 0.08 represents reasonable errors of approximation in the population, and any value greater than 0.10 do not support a fit at all.

GFI (Goodness-of-Fit Index) is a measure of the relative amount of variance and covariance in the model. It ranges from 0 to 1 with values close to 1 being good fit.

NFI (Normed Fit Index) ranges from 0 to 1 on the basis of the comparison of the model with the independence model (null model). A value greater than 0.90 indicates an acceptable fit to the data.

RMR (Root Mean Square Residual) represents the average residual value derived from the fitting of the variance-covariance matrix for the model to the variance-covariance matrix of the sample data. Standardized RMR represents the average value across all standardized residuals. In a well-fitting model the value of a RMR will be small (less than .05).

4.3 Literature review

Overview of various fitness composite scores used in the current literature

To illustrate the use of CMS, published studies in the area of physical fitness measurement were reviewed. Most of the summary approaches do not consider the structure of a measured construct. The first author who considered at least correlations between tests assumed to measure a single construct of functioning was Guralnik (2004). A similar problem was later investigated in three studies exploring dimensionality of physical fitness among older adults. All of them were lead by the same author – H. Nagasaki in 1995 (Kinugasa & Nagasaki, 1998; Nagasaki, Itoh, & Furuna, 1995a, 1995b) and all of them will be described in more detail in the following section. Now, the attention will be focused on other currently published fitness summary scores developed for research purposes. A brief description of different development processes will be provided. This review was also done to emphasize broad use of composite measurement scores in order to point out their evident need.

Recently, the overall fitness score was used by Dobek et al. (Dobek et al., 2006) in a study which examined the effect of a ADL-based exercise training program on performance in both activities of daily living and physical fitness. Authors of this study used the Senior Fitness Test (SFT) tests to compute overall fitness score. The standard scores (Z-scores) from the six tests that create the SFT (30-second stand tests, 30-second arm curl test, 2-minute step test, chair sit-and-reach test, back scratch test, and 8-feet up-and-go test) were averaged to determine the overall change in overall fitness.

Another example was presented by Buchman et al. (Buchman, Boyle et al., 2007). His team tested the hypothesis that physical activity modifies the course of age-related motor decline. Nine strength measures and nine motor performances were summarized into a composite measure of motor function. Composite measure was used in this study because it yields a more stable measure of motor function and increases power to identify the risk factors as well as consequences of motor decline in aging. Muscle strength was measured using portable hand-held dynamometers that are reliable in older persons (Wang, Olson, & Protas, 2002). The hand dynamometer was used to assess muscle strength in both arms (arm abduction, arm flexion, arm extension), and both lower extremities (hip flexion, knee extension, plantar flexion, ankle dorsi flexion). Grip and pinch strength were measured by Jamar hydraulic and pinch dynamometers. The mean score for each muscle group was converted to Z-scores, using the baseline mean and standard deviation of all study participants, which were then averaged to yield a composite measure of strength. Motor performance was tested in both upper and lower extremities. In the lower extremities, the following performance based tests were used – walk 8 feet and turn 360°, stand on each leg for 10-second, stand on the toes for 10-second, walk 8-feet along line in a heel to toe manner. In this case, the factor analysis was performed with a result of two factors: gait and balance. Component measure of gait and balance were then formatted by converting each of the motor tests to a Z-score to form composite gait and balance measures. Two tests of upper-extremity motor performance were used. The number of the peg that could be placed in 30-seconds was recorded and participants tapped and electronic tapper with their index finger as quickly as possible for 10-second. A measure of upper-extremity motor performance was created by converting Purge pegboard and finger tapping scores to a Z-score and then computing the average of the Z-scores. A composite measure of muscle strength was formed by averaging the Z-scores for arm and leg strength. A composite measure of motor performance was made by averaging the Z-score for gait (lower-extremity motor performance), balance and upper-extremity motor performance. A composite measure of global motor function was created by averaging all of the motor function tests together. Buchman and his colleagues used a similar composite measure of motor performance and muscle strength in their study focused on the association between changes in motor function and mortality (Buchman, Wilson, Boyle, Bienias, & Bennett, 2007).

Another approach was made by Spanish researchers who assessed the validity of a battery of functional capacity tests (Avila-Funes, Gray-Donald, & Payette, 2006). The global score measuring functional capacities score was constructed as the sum of four tests. They were motivated by the approach of Guralnik who developed widely used lower-body performance SPPB battery (Guralnik, Simonsick et al., 1994). A summary performance scale was created by summing categorical rankings of performance on each test. This instrument was used in the Established Populations for Epidemiologic Studies of the Elderly (EPESE), a collaborative longitudinal study of aging initiated and funded by the National Institute of Aging. Summing quartile scores for tests of standing balance, gait speed, and rising from a chair five times formed quartile summary physical performance score. These scores have been validated and have a good reliability (Guralnik, Seeman et al., 1994; Guralnik, Simonsick et al., 1994; Ostir, Volpato, Fried, Chaves, & Guralnik, 2002). Also they successfully discriminate between groups in observational studies (Guralnik, Ferrucci, Simonsick, Salive, & Wallace, 1995; Guralnik, Seeman et al., 1994; Ostir et al., 2002; Tinetti, Doucette, Claus, & Marottoli, 1995). However, quartile scores may not be responsive to small differences in physical function within an individual at the lower end of function. Therefore, quartile scores may have limited use in clinical studies of interventions on functional decline in the elderly, in which treatment course is relatively short and changes in functional status are small.

To address the potential limitations of the quartile scoring method, a continuous summary score based on outcome measures of the performance test has been developed. Scores from individual tests can be summed to produce a continuous summary score (Onder et al., 2002). This approach was published by Nieves et al. (Nieves et al., 2005) and the conclusion from this study was that in cohorts with moderate to severe disability, the continuous version of scoring appears to be valid and reproducible measure that can discriminate smaller yet clinically meaningful differences in physical function, as compared to the quartile scoring.

Those are examples of previously published studies using composite measurements scores. Except for Buchman's study where at least factor analysis was performed to test the structure of lower-extremities, nobody considered either unidimensionality of the set of tests or different contribution of included tests to a measured composite score.

Overview of fitness models considering unidimensionality testing

The structure of fitness in young people is viewed as a multidimensional construct, in that it consists of major independent components such as strength, speed, aerobic endurance, and flexibility. If this was a true for older adults, it would be nonsense to even think of any summary score described earlier. But is the structure of fitness in older adults really the same as in young population? We have detected three studies which applied testing of structure model into their designs. Two of them used performance measures and one of them self reports. All three studies analyzed data by CALIS procedure in SAS system.

The first study is called: *A Physical Fitness model of older adults* (Nagasaki et al., 1995a). This study examined whether the structure underlying fitness in young adults was also relevant for older adults. A 10-item performance battery assumed to assess six components of physical fitness was administered to sixty nine healthy volunteers ranging from 61 to 83 years. A covariance structure model was applied to the test data (figure 4.2): the second-order factor was Physical Fitness, and the first-order factors were strength, mobility, balance, flexibility, stamina, and manual speed (1-6 in figure 4.2). All of the first-order factors were measured by ten manifest variables/indicators (I1 - I10 in figure 4.2). Goodness-of-fit index of the model was acceptable (GFI = 0.93). While four factors relating to basic motor performances (strength, mobility, balance, and manual speed) had loadings more than 0.62 to the physical fitness, flexibility and stamina had loadings less than 0.35. It was demonstrated that in older adults strength, mobility, balance, and speed fitness underlie a single common factor, but the flexibility and stamina not.

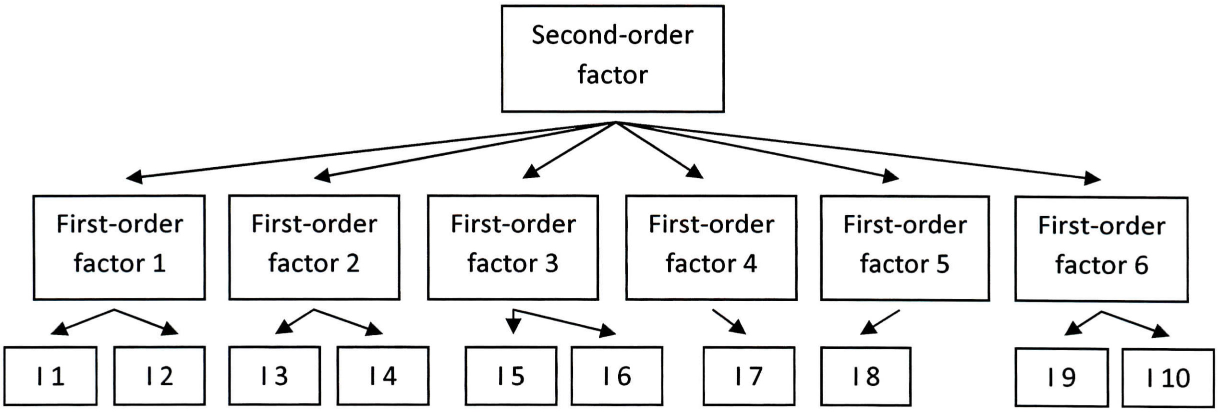


Figure 4.2 Schematic covariance structure model of Physical Fitness developed by Nagasaki et al. (1995)

The second study is called: *The structure underlying physical performance measures for older adults in community* (Nagasaki et al., 1995b). This one examined the structure of a 6-item physical performance battery assessing hand strength/speed, mobility, and balance. The same second-order covariance structure model was applied to the data (figure 4.3). The second-order factor was called Basic Motor Ability (BMA), and the first-order factors were hand strength/speed, walking, and balance (1-3 in figure 4.3). All of them were measured by six manifest variables/indicators (I1 - I6 in figure 4.3). All three factors had loadings more than 0.80 on a single first-order factor. The BMA score was calculated on the basis of this model as a summary score of six physical performances. The BMA predicted the self-reported levels of competence and physical activity with greater accuracy than age alone. The BMA also differentiated those at the high end of the functional spectrum, and thereby not identifiable by use of ordinary self-reported functional measures.

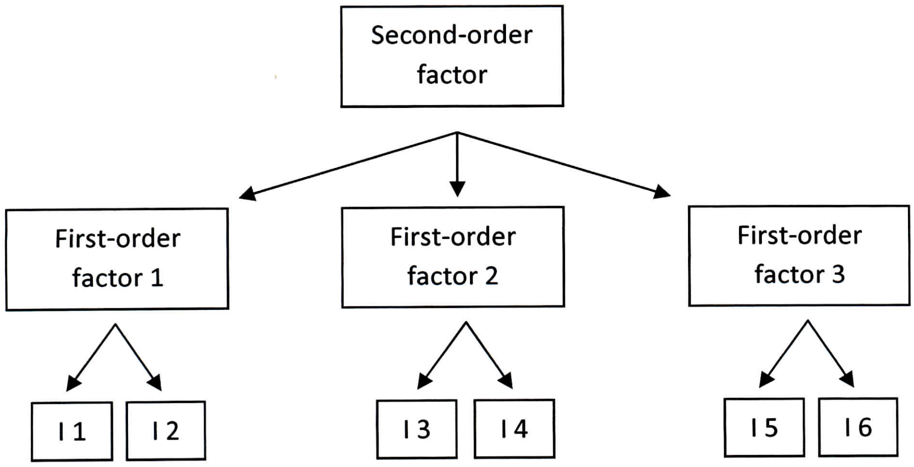


Figure 4.3 Schematic covariance structure of Basic Motor Ability developed by Nagasaki et al. (1995)

Later in 1998, the same approach was applied on the self-report data in the study called: *Reliability and validity of the Motor Fitness Scale for older adults in the community* (Kinugasa & Nagasaki, 1998). The authors constructed structure model of Motor Fitness Scale as a construct consisting of three components: mobility, strength, and balance. The Motor Fitness Scale was a second-order latent factor and the components were first-order latent factors. As indicated, each first-order factor was assumed to have loadings on the second-order factor, Motor Fitness Scale. Each indicator variable assessed by a self-report question was assumed to have loadings on one of the first-order factor. Goodness-of-fit index, adjusted goodness-of-fit index, root mean square residual and χ^2 were used as indices of the model's fit. A unidimensional structure was confirmed. This model accounts for 93% of variances in the data. Similarly as in the previous approaches, the Motor Fitness scale does not cover such components of physical fitness as stamina or flexibility, which are also important for older adults' independent and successful functioning.

4.4 Summary and hypothesis

The results presented in the literature evince that a use of various composite measurement scores is broad in current research. Composite measurement scores are practical and important for both research and clinical purposes. However, a review of current practice in the area of constructing evaluative composite measurement scores showed serious flaws. Most of the published approaches did not consider basic measurement properties, and even more importantly, they did not examine if the structure of a measured construct is unidimensional. Just three articles investigating the structure of physical fitness among older adults were detected. Given a lack of sufficient process regarding adequate methodology in creating composite measurement scores, this study will focus specifically on a theoretical and statistical testing of an Overall Functional Fitness score. It will carry on in Nagasaki and his colleagues' work in order to confirm and extend their findings. The long term goal of this study was to further the understanding of the structure of functional fitness among older adults. Findings from this study will provide a baseline data that will help to accurately evaluate older adults' Overall Functional Fitness status, an important factor of quality of life in old age. Three hypothesis-driven specific aims will be addressed in this study:

Specific aim 1: to provide the theoretical model of the Overall Functional Fitness and operationalize the construct.

Specific aim 2: to empirically examine the theoretical model of Overall Functional Fitness using structural equation modeling. Performance tests previously validated to measure each functional fitness component will serve as manifest variables for the statistical model.

Specific aim 3: to investigate the contribution of individual components to the overall construct.

Hypothesis

H1 The Overall Functional Fitness among older adults is a unidimensional construct consisting of six individual components.

H2a Each component has its unique contribution to the Overall Functional Fitness.

H2b The most important components are those related to mobility.

4.5 Methods

4.5.1 Participants

The study sample included seventy eight older adults (69 women and 9 men; mean age 81.9; SD 9.31). These older adults were mostly low functioning elderly living in residential care facilities (RCFs), a common type of long-term care settings in the Czech Republic as well as in other European countries. Although the functional limitation is not a condition for admission in the Czech Republic (I Holmerova, 2007), residents in RCFs tend to be more functionally impaired than the rest of the Czech elderly population (Kalvach et al., 2004). Participants were recruited from a larger study, 'Dance and Quality of Life' (Vankova et al., 2008) which was conducted in seven different RCFs in the years 2005-2007. More detailed data were collected from residents in three RCFs in urban area starting in October 2006 and ending in June 2007. These detailed data was used to examine the effect of dance based exercise on functional fitness. All participants signed an informed consent form. Ethical Committee and the Internal Review Board at the Internal Grant Agency of the Ministry of Health of the Czech Republic approved the study. For the purposes of the dissertation we use just the pre-tests of recruited participants. From the total of ninety randomly recruited participants, seventy accepted performance assessment and another eight participants asked to participate even they were not chosen for the dance study.

4.5.2 Procedure

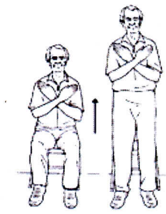
All testing was done in a field setting such as participant's room, activity room, hall, or exercise room. Prior to performing the activities in the SFT battery, all procedures were explained by the administrator and series of warm up activities were performed by the participants. Warm-up activities included marching in place while seated, swinging the arms and stretching large-muscle groups. Participants were asked to do the best they could on all tests but never to push themselves to a point of overexertion or beyond what they thought was safe. The following section will offer a full description of each test within the SFT battery including the equipment requirements, test procedures and scoring instructions (Rikli & Jones, 2001), and the adaptations made in this study.

During testing, participants were closely monitored for signs of overexertion and were discontinued from the study if they were experiencing any of the following symptoms: unusual fatigue or shortness of breath; dizziness; tightness or pain in the chest; irregular heartbeats; pain of any kind; numbness; loss of muscle control or balance; blurred vision; or confusion and disorientation.

4.5.3 Clinical instruments

Senior Fitness Test (Rikli & Jones, 2001)

The SFT was designed to be easily administered without extensive time, equipment, or space requirements. The test battery can be given by a trained instructor in approximately 15 to 20 minutes to an individual or in approximately 30 minutes to a small group (up to five participants). All of the tests were standardized for older population over 60 years of age and satisfactorily reliable. A detail description of each test used in the study is below:

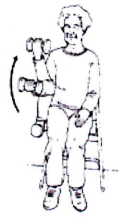

30-second chair stand test

- Equipment: stopwatch and standard chair (height of 43cm) placed against the wall to prevent slipping

- Procedure: the participant was instructed to sit in the middle of the chair with the back straight, feet flat on the floor, and arms crossed at the wrist and held against the chest. On the signal 'start' the participant rose to a full stand, than returned to a fully seated position. The participant was encouraged to complete as many full stands as possible in 30 seconds. The test was demonstrated slowly to illustrate the proper form, than at a faster pace to show that the object is to do the best one can within safety limits. Before testing the participant practiced one or two stands to ensure the proper form.

- Scoring: the score was a total number of stands completed in 30 seconds. Only one test trial was administered.

- Adaptations: if participant could not perform even one stand without using their hands, he or she was allowed to push off by his or her legs or by chair arms (if any). The adaptation was described but the score for the purposes of this study was zero.



30-second arm curl test

- Equipment: stopwatch, standard chair (height of 43cm) and dumbbell (2.27kg for women and 3.63 kg for men)

- Procedure: the participant was in seated position with the back straight and feet flat on the floor, and with the dominant hand close to the edge of the seat. The weight was held down at the side in the dominant hand with a handshake grip. From the down position, the weight was curled up with the palm gradually rotating to a facing-up position during the flexion. The weight was then returned to the fully extended position with the handshake grip. On the signal 'start' the participant curled the weight through the full range of motion as many time as possible in 30 seconds. The upper arm had to remain still during the test. Bracing the elbow against the body helped stabilize the upper arm. The test was demonstrated slowly to illustrate the proper form, then in a faster pace to illustrate the pace. The participant practiced one or two repetitions without and with the weight to ensure the proper form.

- Scoring: the score was a total number of arm curls completed in the 30 seconds. If the arm was more than halfway up at the end of 30 seconds, it counted as a curl. Only one trial was administered.

- Adaptations: slightly lighter dumbbells were used (2.5 kg for women and 3.5 kg for men). The reason was that the original dumbbells were not available in the Czech Republic.


2-minute step test

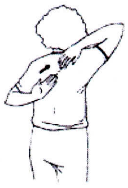
- Equipment: stopwatch, piece of string and masking tape
- Pre-procedure: the minimum stepping height is at the level even with the midway point between the kneecap and the front hip bone. The proper height of steps was determined using a tape measure. The correct height of knee was corrected by moving the participant to the wall or a door way and transferring the tape to a spot at the same level on the wall or between the door.
- Procedure: participant was asked to stand against the wall or door way and on the signal 'start' to begun stepping in a place as many times as possible in the 2-minute period. Although both knees had to be raised to the correct height, only the number of times the right knee reached target was counted. When the proper knee height could not be maintained, the participant was asked to slow down, or to stop until he or she could regain the proper form, but the time was kept running. The test was demonstrated slowly to illustrate the form, then in a faster pace to illustrate the pace. The participant practiced four to six repetitions to ensure proper form.
- Scoring: the score was the number of full steps completed in 2 minutes. Only one trial was administered.
- Adaptations: if participant was unstable, he or she was allowed to hold the wall or chair. If participant was unable to lift their knees to a proper height or could lift only one, he or she was allowed to complete the test, but it was indicated on the score card and zero score was administered for the purposes of final analysis.

 Chair sit-and-reach test

- Equipment: standard chair and ruler

- Procedure: participant was seated on the edge of the chair. He or she extended one leg as straight as possible in front of the hip with the heel on the floor and foot flexed at approximately 90 degrees. The other leg was bent with the foot flat on the floor. With arms outstretch, hands overlapping, and middle fingers even, the participant slowly bended forward at the hip joint reaching as far as possible towards or past the toes. If the extended knee started to bend, the participant was asked to move slowly back until it was straight. The maximum reach must have been held for two seconds. Only the preferred leg was used for scoring purposes. The test was demonstrated and participant practiced one or two times to ensure proper form.

- Scoring: the distance from the tip of the middle fingers to the top of the shoe was measured to the nearest half centimeter. Minus points were given if the middle fingers did not touch the top of the shoe, zero score if the middle fingers just touched, and plus scores if the middle fingers overlapped. Better of two scores was administered.



Back scratch test

- Equipment: ruler
- Procedure: participant was asked to stand and place his or her preferred hand over the same shoulder, palm down and fingers extended, reaching down the middle of the back as far as possible. Then the participant was instructed to place other hand around the back of the waist with palm up, reaching up of the middle of the back as far as possible in an attempt to touch or overlap the extended middle fingers of both hands. The test was demonstrated and participant practiced one or two times to ensure proper form.
- Scoring: the distance between the tips of both middle fingers was measured. Minus points were given if the middle fingers did not touch, zero score if the middle fingers just touched, and plus scores if the fingers overlapped. Better of two scores was administered.

8-foot up-and-go test

- Equipment: stopwatch, standard chair, tape measure and cone
- Setup: the chair was placed by the wall if possible facing a cone marker exactly 2.44 meters (8 feet) away.
- Procedure: participant was instructed to sit in the middle of the chair with back straight, feet flat on the floor, and hands on the thighs. One foot should have

been slightly in front of the other foot, with the torso slightly leaning forward. On the signal 'start' the participant got up from the chair, walked as quickly as possible around either side of the cone, and then sit back down in the chair. The test was demonstrated and participant practiced one time to ensure the proper form.

- Scoring: the trial was administered to the nearest second.
- Adaptations: if needed, a cane or a walker could have been used for this test but the score was not considered for the final analysis.

4.5.4 Data analysis

Data were entered in the SPSS 11.5 program for Widows (SPSS Inc., Chicago, IL, USA) which was used for data processing, descriptive statistics of all variables including correlation matrix necessary for further analysis in Lisrel, and factor analysis. The model shown in figure 4.5 was empirically tested using a correlation matrix of manifest variables obtained from the SFT test battery. The testing was performed using structural equation model in Lisrel 8.8. Whereas traditional multivariate procedures are incapable of either assessing or correcting for measurement error, SEM provides explicit estimates of these parameters. Also, the former methods are based on observed measurements only, SEM can incorporate both unobserved and observed variables (Bollen, 1989). Observed or manifest variables serve as indicators of the underlying construct (latent variable or factor) that they are presumed to represent (figure 4.5). Root mean square error of approximation (RMSEA), normed fit index (NFI), goodness of fit index (GFI), and root mean square residual (RMR) were used as indices of the fit of the model. Also, as a part of LISREL outcomes, we obtained Factor Scores Regressions which determined the contribution of each test within the SFT test battery to its common factor called Overall Functional Fitness.

Comparisons with Other Analysis:

- Path analysis, factor analysis, and regression are special cases of SEM
- SEM is basically path analysis with latent variables
- Path analysis contains only observed variable, and has more restrictive set of assumptions than SEM

Data adjustment method

Findings from the previous chapter showed that the SFT might be sometimes difficult to perform, especially for the oldest participants living in Residential Care Facilities. Some of them were not able to perform task without help of their hand or other person (table 4.1). Few participants were unable to walk longer distance without support or they failed in other SFT items. Importantly, only reason for missing values was inability to perform the task because of the low fitness level.

Table 4.1 Descriptive statistics (minimum, maximum, mean, standard deviation) for individual Senior Fitness Test tests for the study sample

	N (78 total)	Min	Max	Mean	SD
Test 1: 30-second chair stand test [number of repetitions]	73	2	21	8.21	3.67
Test 2: 30-second arm curl test [number of repetitions]	77	3	26	12.60	4.90
Test 3: 2-minute step test [number of repetitions]	63	2	115	35.16	21.33
Test 4: chair sit-and-reach test [cm]	72	-35	20	-6.17	11.51
Test 5: back scratch test [cm]	70	-65	4	-21.79	15.57
Test 6: 8-foot up-and-go test [seconds]	61	5.1	145	20.80	23.37

For the purposes of analysis performed in this study, the priority was to reflect the true fitness status as accurately as possible. To fulfill this requirement, the missing data had to be somehow adjusted. If a participant was not able to perform properly test 1-3, his or her score was administered as 0. This method was sufficient for the half of the SFT tests where 0 meant the worst possible score. But it appeared more complicated for flexibility and agility tests. For instance, the flexibility score 0 would mean that participant just touched the fingers behind his or her back in the upper-body flexibility test or touched the toe by the middle fingers in the lower-body flexibility test. The agility score 0 indicating time to walk would mean nonrealistic score either. Participants who were not able to perform either flexibility or ability tasks should have gained the number that would represent the lowest possible score for each test for the study sample. Thus the following method has been applied: the number representing the worst possible score for both flexibility and ability tests was calculated as the average of the three worst scores in each test. The following numbers were obtained: test 4 = -35 cm, test 5 = -52.5 cm, test 6 = 91 seconds. Table 4.2 presents descriptive statistic for the sample after data adjustment.

Table 4.2 Descriptive statistics (minimum, maximum, mean, standard deviation) for individual Senior Fitness Test tests for the study sample with adjusted missing values

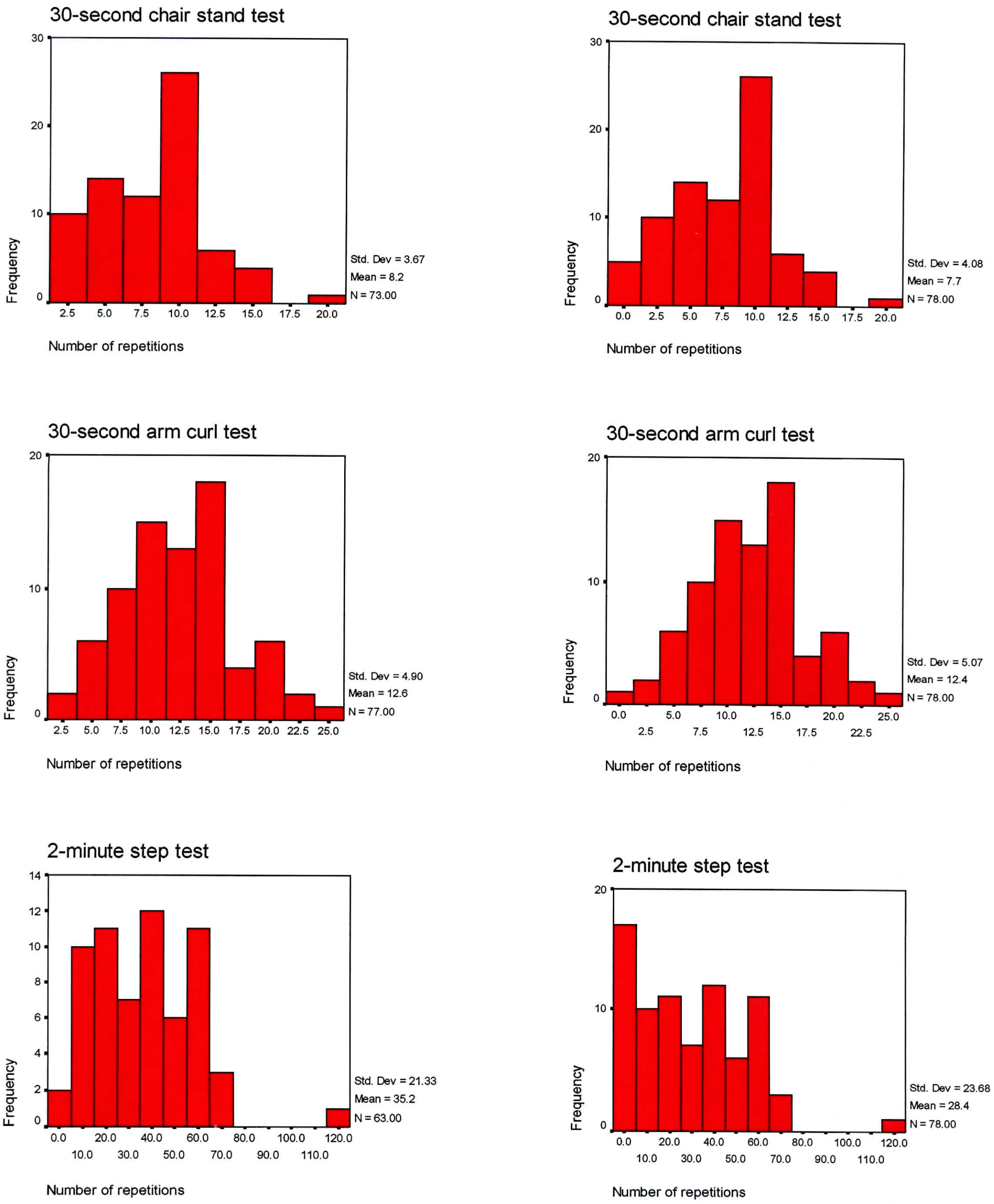
	N (78 total)	**Min**	**Max**	**Mean**	**SD**
Test 1: 30-second chair stand test [number of repetitions]	78	0	21	7.68	4.10
Test 2: 30-second arm curl test [number of repetitions]	78	0	26	12.44	5.10
Test 3: 2-minute step test [number of repetitions]	78	0	115	28.40	23.70
Test 4: chair sit-and-reach test [cm]	78	-35	20	- 8.4	13.50
Test 5: back scratch test [cm]	78	-65	4	-25.85	19.00
Test 6: 8-foot up-and-go test [seconds]	78	5.1	145	39.3	41.00

To assess the legitimacy of our approach we compared a data distribution (skewness and kurtosis) of the sample before and after data adjustment (table 4.3). Even though, the distribution was not perfect either before-adjustment or after-adjustment, the latter data set was closer to the normal distribution for most of tests.

Table 4.3 Data distribution (skewness and kurtosis) for individual Senior Fitness Test tests before and after missing data adjustment

	Skewness Non-adjusted data	**Kurtosis** Non-adjusted data	**Skewness** Adjusted data	**Kurtosis** Adjusted data
Test 1: 30-second chair stand test [number of repetitions]	0.434	0.671	0.247	0.332
Test 2: 30-second arm curl test [number of repetitions]	0.074	-0.359	-0.075	-0.153
Test 3: 2-minute step test [number of repetitions]	0.784	1.189	0.801	0.918
Test 4: chair sit-and-reach test [cm]	0.663	-0.186	-0.666	-0.325
Test 5: back scratch test [cm]	0.398	0.160	-0.455	-0.065
Test 6: 8-foot up-and-go test [seconds]	4.022	20.465	1.280	0.401

Skewness is a measure of symmetry, or more precisely, a lack of symmetry. A distribution of data set is symmetric if it looks the same to the left and right of the center point. Kurtosis is a measure of whether the data are peaked or flat relative to the normal distribution. That is data sets with high kurtosis tend to have a distinct peak near the mean, decline rather rapidly, and have heavy tails. Data sets with low kurtosis tend to have a flat top near the mean rather than a sharp peak. A uniform distribution would be the extreme case. The histogram is an effective graphical technique for showing both the skewness and kurtosis of data set. Negative values for the skewness indicate data that are skewed left and positive values for the skewness indicate data that are skewed right. The kurtosis for a standard normal distribution is three. For this reason, some sources use adjusted definition of kurtosis so that the standard normal distribution has a kurtosis of zero. In addition, with the adjusted definition positive kurtosis indicates a "peaked" distribution and negative kurtosis indicates a "flat" distribution. Which definition of kurtosis is used is a matter of convention. When using software to compute the sample kurtosis, you need to be aware of which convention is being followed. In this study, the adjusted kurtosis definition was used. Histograms for all tests for both datasets (before and after missing data adjustment) are show in the following figure 4.4.



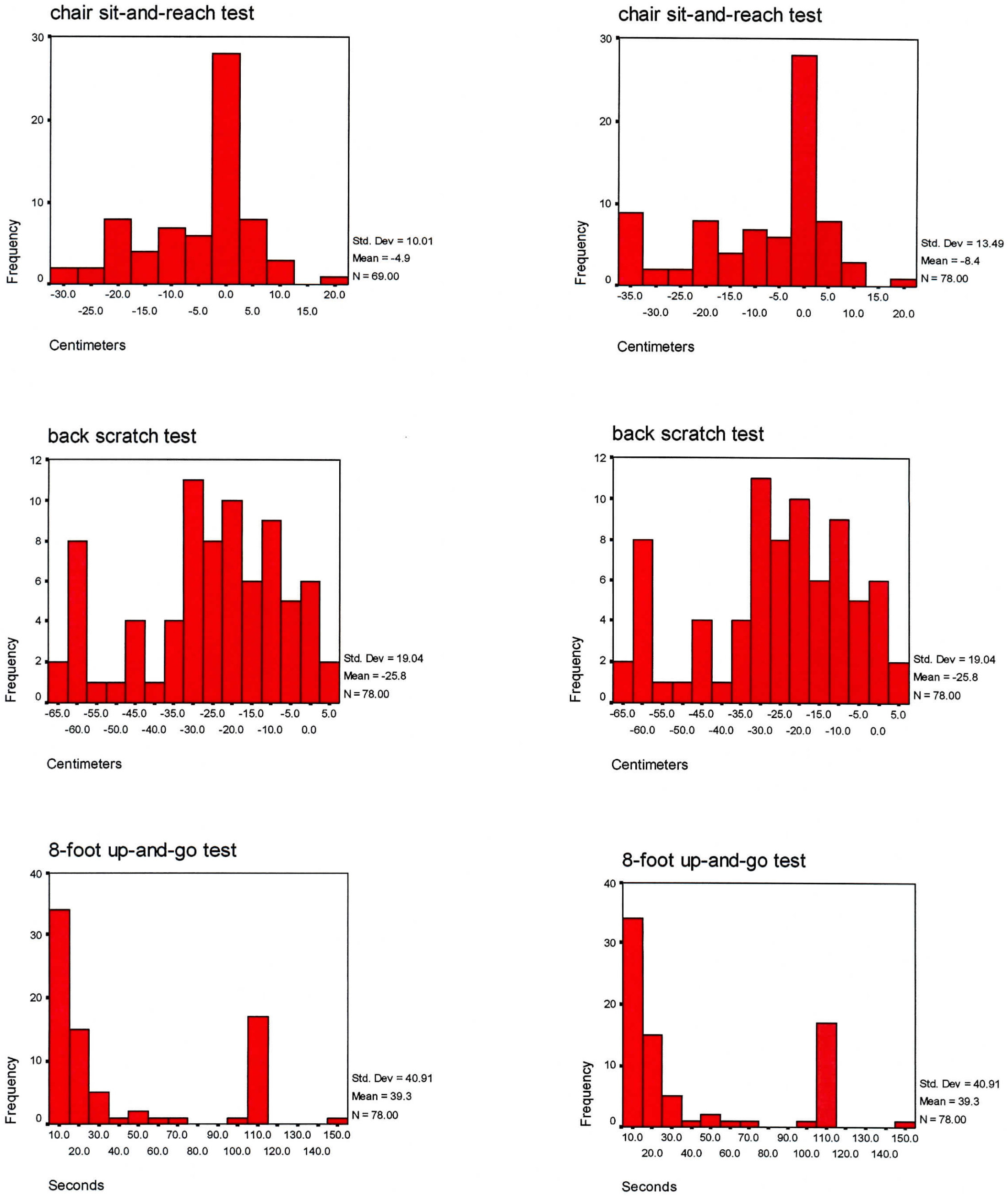
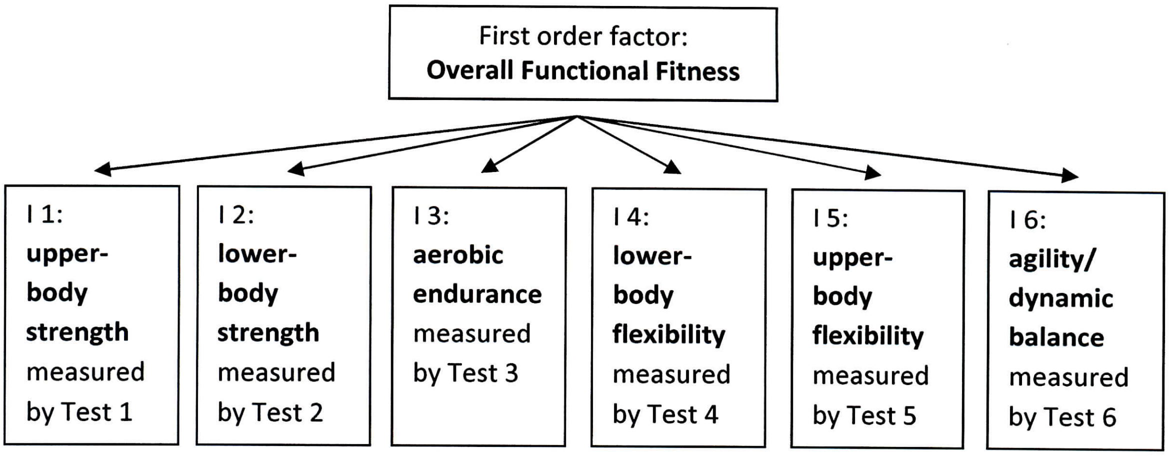


Figure 4.4 **Histograms illustrating distribution of individual Senior Fitness Test tests before (left column) and after (right column) missing values adjustment**

4.6 Results

Theoretical measurement modeling

The common activities are taking care of personal needs, household chores, shopping, or traveling. These require the ability to perform functions such as walking, stair climbing, lifting, and reaching and these function, in turn require an adequate reserve in the physical fitness components (table 1.1). Based on this framework and the evidence described in the general introduction, following components were identified as relevant for functional fitness: muscle strength, aerobic endurance, flexibility, agility/dynamic balance (Frankel et al., 2006; Rikli & Jones, 2001; Shepard, 1997; Spirduso, 2005; Toraman, 2005). According to the theory, a structural model of the Overall Functional Fitness constructed as illustrated in figure 4.5 was applied to the results of performance test. The model was defined as a first-order covariance structure in which the Overall Functional Fitness was the only latent variable – a first-order factor. Six fitness components were represented by six performance tests which were previously validated to measure corresponding components. Test 1 and test 2 measured upper- and lower-body muscle strength, test 3 measured aerobic endurance, test 4 and test 5 measured lower- and upper-body flexibility and test 6 measured agility/dynamic balance.



Note: Test 1 – 30-second chair stand test
Test 2 – 30-second arm curl test
Test 3 – 2-minute step test
Test 4 – Chair sit-and-reach test
Test 5 – Back scratch test
Test 6 – 8-foot up-and-go test

Figure 4.5 Schematic covariance structure model of Overall Functional Fitness

To visually examine the structure of our data we evaluated the correlation matrix (table 4.4). The correlation matrix is one of the initial steps needed for further approaches such as SEM. The adjusted dataset of seventy eight participants was used because it was considered to reflect reality more accurately as compared to the original dataset.

Table 4.4 Correlation matrix of individual Senior Fitness Test tests

	Test 1	**Test 2**	**Test 3**	**Test 4**	**Test 5**	**Test 6**
Test 1	1	-	-	-	-	-
Test 2	0.543**	1	-	-	-	-
Test 3	0.595**	0.426**	1	-	-	-
Test 4	0.533**	0.431**	0.454**	1	-	-
Test 5	0.356**	0.337**	0.399**	0.399**	1	-
Test 6	-0.511**	-0.325**	-0.514**	-0.481**	0.458**	1

Note: **p<.01

Also before actual model testing in Lisrel, the structure was pre-determined using factor analysis. It appeared that single factor explained 54.480 % of total variance (table 4.5). In addition, the measure of internal consistency, Cronbach's alpha, was computed. The presented evidence including: the correlation matrix; the factor analysis; and the internal consistency coefficient (α = 0.8312) was considered sufficient and supportive for the composite measurement score testing in Lisrel.

Table 4.5 Percentage of variance explained by a single common factor – results of factor analysis performed in SPSS

Component	**% of variance explained**
1	54.480
2	12.641
3	10.426
4	9.140
5	7.364
6	5.949

Statistical measurement modeling

Based on the theoretical modeling and pre-analysis presented above, the model of Overall Functional Fitness was finally tested using structural equation modeling performed in Lisrel. It was assumed that the Overall Functional Fitness consists of six major components: muscular strength – lower-body and upper-body; aerobic endurance; flexibility – lower-body and upper-body; and agility/dynamic balance. Each of them was measured by single standardized performance test. The model in figure 4.5 illustrates simple structure where the Overall Functional Fitness represented a first-order latent variable. Indicators (manifest variables) were represented by six standardized tests from the Senior Fitness Test battery (Rikli & Jones, 2001). Each of them was validated to measure individual components of functional fitness. Basically, The SFT tests identified the expected model path. Each test/indicator was assumed to have loadings on the first-order factor, in this case on the Overall Functional Fitness. Except for the goodness-of-fit indices, the factor score regressions were obtained. If model fits, factor score regressions (table 4.6) will represent contributions of each test to the Overall Functional Fitness. The result of the structural equation model is shown in figure 4.6.

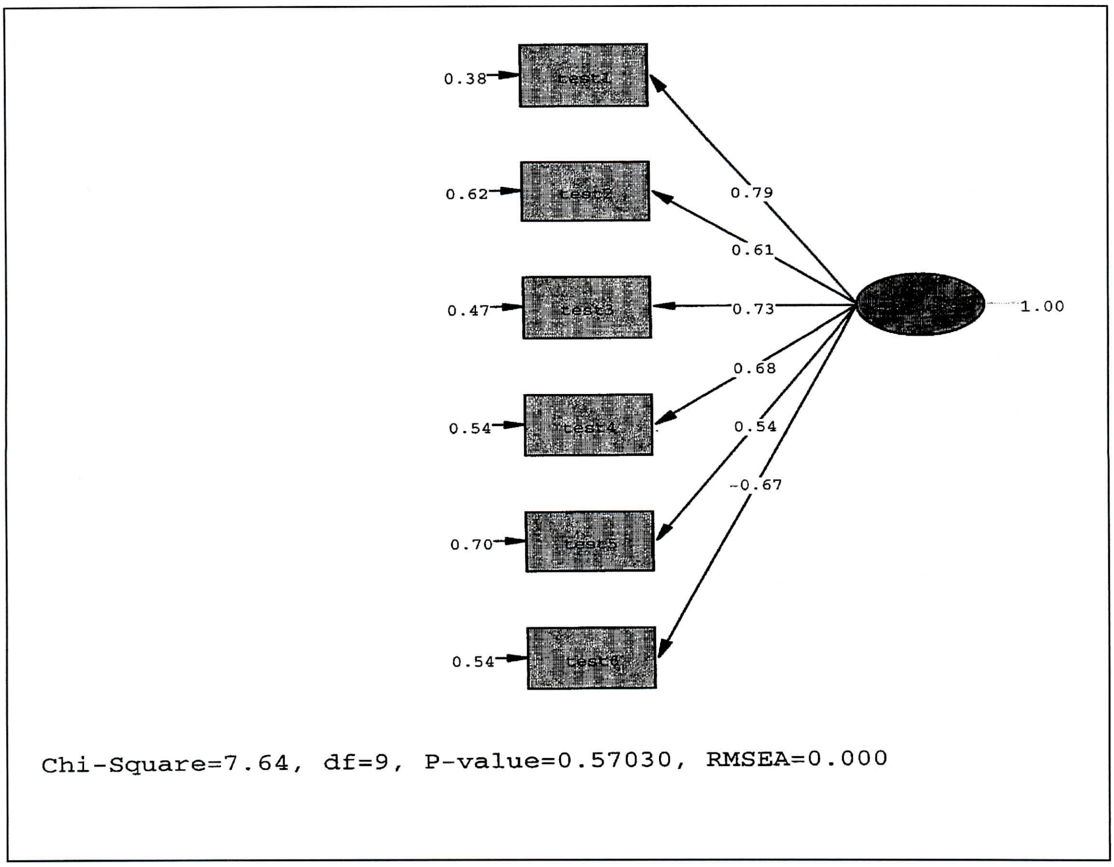


Figure 4.6 Schema of standardized maximum likelihood solution for the model of Overall Functional Fitness (N = 78; RMSEA = .00; NFI - .95; GFI = .97; RMR = 0.038; $\chi2$ = 7.64 (df = 9))

Goodness-of-fit indices

Chi-square = 7.64 (*p* = .57)

Degrees of freedom = 9

Root mean Square error of approximation (RMSEA) = 0.00

90% Confidence interval for RMSEA = (0.0; 0.11)

Goodness of Fit Index (GFI) = 0.97

Root Mean Square residuals (RMR) = 0.038

Table 4.6 Factor scores regressions representing contribution of individual fitness components to Overall Functional Fitness – result of structural equation model performed in Lisrel

	Test 1	Test 2	Test 3	Test 4	Test 5	Test 6
Functional fitness indicators	0.32	0.15	0.24	0.19	0.12	-0.19

Goodness-of-fit indices indicated that the theoretical model fits very well. Original Lisrel output is attached as appendix A1. Because the fit of the model was satisfactory, the factor scores regressions represent the contribution of each test to the latent construct and may be used as weights when computing composite score. Specifically, the findings revealed hierarchical structure of functional fitness. The most important component was lower-body muscle strength measured by the 30-second chair stand test – test 1 (0.32) followed by aerobic endurance measured by the 2-minute step test – test 3 (0.24), agility measured by the 8-food up-and-go test – test 6 (-0.19) and lower-body flexibility measured by the chair sit-and-reach test – test 4 (0.19). Even though upper-body functioning might seem minor, it is also very important component within the construct and cannot be eliminated. Upper body strength measured by the 30-second arm curl test – test 2 gained the factor score of 0.15 and upper-body performance measured by the back scratch test – test 5 gained the factor score of 0.12. Based on the results, the most accurate estimation of the Overall Functional Fitness would be a weighted sum of each fitness component measured by relevant manifest variable. So in this case, it would be a weighted sum of the 30-second chair stand test, 30-second arm-curl test, 2-minute step test, chair sit-and-reach test, back-scratch test and 8-food up-and-go test.

Application of the results into practice

Because all tests were administered in different units such as: a number of repetitions (test 1 – test 3); centimeters (test 4 and test 5); or seconds (test 6), the data had to be transferred into standardized scores – Z-scores, using the baseline mean and standard deviation of all participants from the study. The equation for computing Overall Functional Fitness would be as follow:

Overall Functional Fitness score = test 1 x 0.32 + test 2 x 0.15 + test 3 x 0.24 + test 4 x 0.19 + test 5 x 0.12 + test 6 x (-0.19)

To illustrate the application of results into practice, two extreme subjects were selected. Subject # 1 was considered as a low functioning and conversely, subject # 2 was considered as a high functioning. The original scores of both subjects for each test are shown in table 4.7 as well as scores representing the Overall Functional Fitness. As can be seen in table 4.7, the Overall Functional Fitness of subject # 1 was 1.327 SD below the sample mean while the Overall Functional Fitness of subject # 2 was 1.457 SD above the sample mean.

Table 4.7 Overall Functional Fitness score of two extreme subjects: illustration of the process of composite score development according to the results from original data, through Z-score transformation, to weighted Z-scores and final summing

	Subject #1			**Subject # 2**		
SFT	**Original scores**	**Z-scores**	**Weighted Z-scores**	**Original scores**	**Z-scores**	**Weighted Z-scores**
Test 1	2 times	-1.391	-0.445	11 times	0.813	0.260
Test 2	10 times	-0.480	-0.072	26 times	2.674	0.401
Test 3	0 times	-1.199	-0.288	60 times	1.334	0.320
Test 4	-23 cm	-1.083	-0.206	10 cm	1.363	0.259
Test 5	-27 cm	-0.061	-0.007	-7 cm	0.990	0.119
Test 6	105.5 sec	-1.625	-0.309	18.3 sec	0.514	0.098
Overall Functional Fitness			**-1.327 SD**			**1.457 SD**

To support meaningfulness of this approach, the sample was divided into three different age categories. We used age categories presented by Holmerova at al. (I. Holmerova, Juraskova, & Zikmundova, 2003). The most numerous group was the middle category:

75 - 89 years of age which consisted of forty nine women. Both of side categories (the youngest and the oldest) consisted of ten women (table 4.8). Despite of unequal group sizes, it is clearly seen that the Overall Functional Fitness deteriorates with age in women. Unfortunately, this trend has not been supported for men most likely because the sample size which was too low to evaluate the effect of age. The results presented in table 4.9 have no research or statistical value. Figure 4.6 just visualized the function of age on the Overall Function Fitness. The ANOVA was performed to statistically compare means suggesting that the means for each age category among women are significantly different (F = 5.934; sig. = .004).

Table 4.8 Descriptive statistics (minimum, maximum, mean, standard deviation) for Overall Functional Fitness for different age categories among women

Age [years]	**N = 69**	**Min [SD]**	**Max [SD]**	**Mean [SD]**	**SD**
60 – 74	10	-0.671	1.721	0.539	0.753
75 – 89	49	-1.928	1.378	-0.093	0.816
90 and over	10	-1.869	1.042	-0.745	0.988

Table 4.9 Descriptive statistics (minimum, maximum, mean, standard deviation) for Overall Functional Fitness for different age categories among men

Age [years]	**N = 9**	**Min [SD]**	**Max [SD]**	**Mean [SD]**	**SD**
60 – 74	2	-0.847	0.705	-0.071	1.097
75 – 89	7	0.279	2.097	0.966	0.581

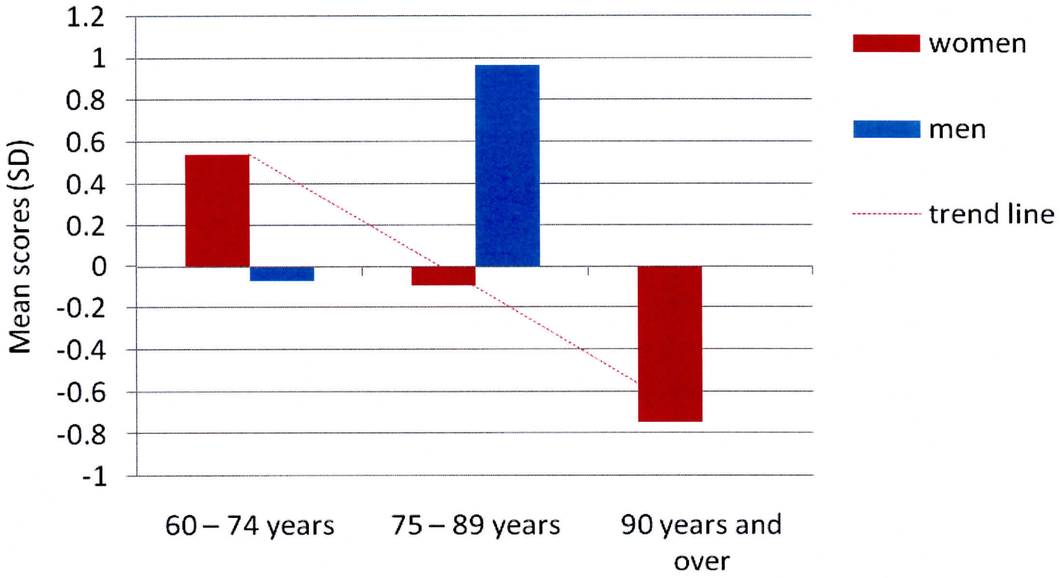


Figure 4.7 Mean Overall Functional Fitness scores for both women (red) and men (blue) for different age categories

4.7 Discussion

This study was proposed as a framework to help investigators in their efforts to estimate functional fitness in older adults more accurately and with the confidence that the summary score is reasonable. But the process as itself applied in this study may be used as an inspiration and guide for any other composite measurement score development in relevant areas of behavioral research and beyond.

The results of demonstrated that the Overall Functional Fitness among older adults is a unidimensional construct. The theoretical model (figure 4.5) which has been developed was confirmed and there was no need for further re-specifications. The goodness-of-fit indices provided us with the sufficient evidence for the construct validity thus the **hypothesis H1 was accepted**. The presented study makes it reasonable to create a single composite score estimating functional fitness of older individuals.

According to the findings, functional fitness consists of six essential components: lower-body strength, upper-body strength, aerobic endurance, lower-body flexibility, upper-body flexibility, and dynamic balance. This results are similar to those provided by Guralnik (Guralnik, Seeman et al., 1994) who regarded six domains of physical performance: balance; gait; upper and lower body strength; hand skills; and coordination as belongings to a single factor. Similar findings were reported by Nagasaki et al. (Nagasaki et al., 1995a, 1995b) but different structure was applied. Nagasaki used two levels structure models in both of his studies. The findings from the first study (figure 4.3) revealed that only four domains: strength; walking; balance; and manual speed out of six can be explained by a single factor. Flexibility and aerobic endurance cannot. His second study (figure 4.4) confirmed unidimensional structure of Basic Motor Ability which consisted of hand power, walking, and balance. Despite of the Nagasaki's two levels models, we decided to test just a single level one. It has been assumed that each component may be measured by a single indicator. This translated into a link between variables representing earlier mentioned components and functional fitness (figure 4.5). To fulfill the theory of functional fitness among older adults and the requirement of multiple testing of each component, the measurement instrument would have to include twelve to eighteen performance tests (two or three tests for each component) which would be too exhausting for the majority of older adults. Our priority was to test rather complete model by single measures than just part of it by multiple measures. Therefore, the two levels structure was rejected.

In addition to the previous research, the contribution of each functional fitness component was investigated. To our best knowledge, this approach has not been studied yet. It is logical to expect that some components are more important compared to another. For instance, it may be assumed that the lower-body functioning which was demonstrated to be essential for mobility (Guralnik et al., 2000; Guralnik et al., 1993) may be more important than upper-body flexibility. According to our findings, all of the components are essential for independent functioning but each contributes differently. As hypothesized, factor scores regressions (table 4.6) clearly demonstrated the hierarchical structure so that measures relevant to mobility and lower-body functioning are the most important for the Overall Functional Fitness while upper-body flexibility is the least important. According to the results, both of the hypotheses 2, specifically, **hypothesis H2a and hypothesis H2b were accepted**.

Based on this specific finding, it appeared that more accurate approximation of functional fitness would be a weighted sum rather than just a simple sum of standardized scores. This may not be as important for descriptive research where a baseline data are the foci of an attention, but it is absolutely essential for experimental research where the effect of intervention is examined. For instance Dobek (Dobek et al., 2006) in his study tested an effect of novel exercise program on functional fitness measured as well as in this study by the Senior Fitness Test. The effect was analyzed for both individual tests of the SFT battery and a composite score. He found statistically significant changes in the tests with the highest importance and no statistically significant changes in the overall score which was created as a simple sum of standardized scores. It is quite possible that if the weighted sum was applied, the improvement in the overall score would become significant because the most important component, which also changed the most, would be advantaged.

Although when creating any composite score, a part of an original information is lost in both cases, the weighted sum have a power to emphasize the most important components which makes the approximation more accurate and also more of the original information is remained. For instance, disabled person on a wheelchair with very strong upper-body performance will not receive higher fitness score compared to other person who is able to perform independently all tests but on lower levels. It is important that despite of the partial information loss, the Overall Functional Fitness remained the ability to distinguish between three different age groups as presented in figure 4.7.

Certain constraints apply to the interpretation of our findings. Firstly, the most accurate composite measure would consist of the items which are considered to be a 'gold standard' or each component of the composite measure would be measured by multiple tests as by Nagasaki. For instance the cardiovascular endurance could be measured by half-mile walk, 2-minute step test, and 6-minute walk test or by maximum VO^2max examined on the treadmill. Unfortunately, even though this practice would desirable from the statistical point of view but is not realistic with the respect to older individuals, their capacities, and early fatigue. Secondly, our sample was relatively small although even though smaller sample sizes using SEM were previously published (Nagasaki et al., 1995a). Thirdly, the participants were lower functioning than general older population might be. It is necessary to replicate this study on a higher functioning sample as well to support the results. And finally, data for most of the tests were slightly skewed and thus did not meet the normal distribution standards although, the Overall Functional Fitness score was distribute quite normally.

CHAPTER 5 – General discussion and conclusions

Evaluation of physical functioning plays a valuable role in clinical geriatrics as well as in aging research. The present dissertation was conducted at the first place to fill the gap and gain an experience with objective measures of functional fitness in older adults living in the Czech Republic. The first problem to deal with was that in the Czech Republic, there was no instrument standardized to measure fitness components in aged population available. It is important to keep in mind that instruments developed for younger age groups are inappropriate. They are usually too demanding, unsafe, and quite often they are even focused on components that are irrelevant for independent living such may be for example speed. Therefore, already standardized instrument from abroad was transferred. The main reasons were following. To develop a brand new instrument is quite complicated process and requires a lot of effort from both the researchers' and subjects' sides. Also the use of already existing instrument provides many advantages such as comparability across cultures and accumulation of findings. Furthermore, there has been developed and standardized many different instrument (see chapter 2) so to create new one would not make much sense.

At the end of 20th century physical functioning has been assessed through self-reports. However, phrasing and responses to these instruments have not been uniform. Concerns about reproducibility, ability to capture the spectrum of disability, precision, and sensitivity to a change have led to a development of performance based instruments (Guralnik, Seeman et al., 1994). These instruments measure functioning quantitatively, often by timing or by applying objective scoring rules such as counting the number of repetition, distance walked, or flights of stairs climbed. It has been documented that performance-based measures earlier identify more limitations compared to self report (Brach et al., 2002; Myers, Holliday, Harvey, & Hutchinson, 1993; Sherman & Reuben, 1998) but the choice of which measurement to use in assessing physical functioning should be based on the research objective, the population under study and many other relevant factors (Kivinen, Sulkava, Halonen, & Nissinen, 1998). It seems that self-report and performance instruments do not measure exactly the same construct (Reuben et al., 1995) in that the performance tests examine the impairment, whereas self-reports reflects disability (Kivinen et al., 1998).

Because there are so many different instruments developed to evaluate physical functioning, the first general aim of this dissertation was to organize available methods and provide the most commonly used examples for each of the methods. This problematic was addressed in the chapter 2 where available instruments were summarized and described. Instruments were organized by the method of testing. Two broad categories were recognized – subjective methods based on self-reports and objective methods based on performance tests. The latter category was further divided into two sub-categories. The first one quantified performance of actual daily tasks and the second one quantified capacity of the organism necessary to perform daily tasks. Again, the latter sub-category was divided into laboratory tests and field tests which were finally split into individual tests and test batteries (figure 2.1). This extensive review of available instruments helped us to select the most appropriate instrument for the present dissertation but most importantly, this review might be a very helpful and valuable source of information for anybody who is interested in fitness assessment in aged population. In addition, the most important issues to consider when selecting appropriate instrument were in detail discussed as well. Based on objectives of our study, it appeared that Senior Fitness Test (Rikli & Jones, 2001) is the most appropriate one.

Therefore, the next chapter addressed the second general aim which was targeted on an application of the SFT battery in Czech older population. Specific aims and hypothesis were empirically tested and in detail discussed in the chapter 3. Hence, just a brief summary will be presented in this section. To make sure that the SFT will be accepted by the majority of older population, two groups of older adults form completely different backgrounds were recruited. The priority was to test both higher and lower functioning individuals. Higher functioning sample served as a pilot testing for this dissertation. As was expected, nobody experience any difficulty with any of tests and all participants accepted performance testing very well. But more important for us was the reaction of the low functioning sample. Even though, some participants, especially those in advanced old age, experienced difficulty with performing some of the tests, in general everybody accepted performance testing surprisingly well. It seemed that most of older adults living in Residential Care Facilities appreciated the attention and even the challenge of testing. Many of them were competitive so they motivated those who were not very confident at the beginning. Also, many older adults realized that they were still able to accomplish some of the tasks which they already believed that are impossible. Further, because it has been proven, that functional fitness may be improved by physical activity even among low functioning older adults, it is crucial to be able to assess baseline levels by performance tests rather that by self-reports, which are not sensitive enough to a change. So even some of the low functioning older adults would score zero before potential intervention, it is very likely that after the intervention they would score 3. For instance, in the 30-second chair stand test, the change from 0 to 3 represents a great accomplishment and may be clinically more important that improvement from 7 to 10. The application of the performance testing even in Residential Care Facilities was considered very successful and promising for the future. Although the SFT was developed on community dwelling older adults this battery seems to very useful for testing lower functioning older adults as it has been previously documented by Beck (2008). It seems that this battery suffer from minimum ceiling and floor effects. The preliminary results of this study has been already published (Machacova, Bunc, Vankova, Holmerova, & Veleta, 2007). In addition, this study may serve as pilot testing for future potential study conducted to develop normative tables.

Finally, the last study was focused on a theoretical approach of development of composite measurement score of functional fitness. Physicians and medical researchers are faced with the necessity of measuring complex constructs such as disease risk or severity, physical disability, functional fitness, or quality of life. The development and psychometric evaluation of scales which measure unobservable - latent constructs continues to be an issue of high interest among many researchers. Meaningful summary measures are needed because the use of multiple evaluations is not always satisfactory or even desirable. Functional fitness is a latent construct consisting of specific components essential for independent functioning (muscle strength, aerobic endurance, flexibility, and agility) so the psychometric evaluation should follow specific procedures (figure 4.1) including both theoretical and statistical approaches. Even there has been published many different composite measurement scores of physical functioning (Avila-Funes et al., 2006; Buchman, Boyle et al., 2007; Buchman, Wilson et al., 2007; Dobek et al., 2006) most of them except for three studies (Kinugasa & Nagasaki, 1998; Nagasaki et al., 1995a, 1995b) did not consider examination of the structure of a measured construct which is essential prior any meaningful composite score development. According to the findings of this chapter, the structure of functional fitness is unidimensional which in other words means that all tests measure a single underlying common factor and that the Overall Functional Fitness score is meaningful and may be constructed. Our findings supported to some point the previous results presented by Nagasaki and his colleagues and extended their work by analyzing the contribution of individual components to the overall construct. As expected, it appeared that each component of functional fitness has different contribution to the overall score and that the structure is hierarchical. The most important components are those related to mobility as proposed by Guralnik (2000) and the least, but also important are upper-body performances.

1. Lower-body strength

    2. Aerobic endurance

        3. - 4. Agility and Lower-body flexibility

            5. Upper-body strength

                6. Upper body flexibility

The findings of this last study are crucial in order to evaluate the Overall Functional Fitness more accurately compared to a summary score proposed by Dobek (2006) and his colleagues. But even more importantly, they will help to create more effective intervention programs. It is evident that interventions created to maintain independent functioning should include stressing all of the components needed for performing daily activities with the special attention to the lower body functioning and aerobic endurance because those appeared the most important for self-supported life. In addition, this study may be use as a guide through the process of composite score development in any other area of behavioral research and beyond.

At the end, three possible ways of interpretation of the data are proposed. Each of the interpretations may be useful for different purposes or needs. Firstly, results may be interpreted in original units. Original scores contain full information and are very important for research purposes and for those who are closely interested in physical functioning and those who need as accurate data as possible. Therefore this way of interpretation would be appropriate for researchers, physical therapists, or even for conditional trainers.

But the information about fitness status may be also highly relevant for others, who are working with older adults on daily basis but might not be as familiar with physical functioning as physical specialists thus they might not know what original scores actually mean from practical point of view. For instance, a nurse in a Residential Care Facility or a caregiver would find helpful and useful to know what to expect from patients but the score of 10 in test 1 would not necessarily provide them with a proper idea of what fitness level does number 10 actually reflects. Therefore, an alternative method of data collapsed into categories may be more appropriate. The easiest way would be to use quartiles or any other percentiles depending on a sample data distribution. Then, the score of 10 in the test 1 could be translated into a category 3 on a scale from 0 to 4 where the score 4 means the highest possible result and the score 0 means that the participant was not able to perform the test. This second way of interpretation would provide more accurate idea about the actual capabilities of participants for any professionals working with older adults. Categories may be also valuable for low functioning older adults who are expected to have some difficulty with performing some of SFT tests. Under some circumstances, the qualitative data may be accompanied by the qualitative description of capabilities.

The same process may be applied on the Overall Functional Fitness score. But one condition should be remained. According to the results, the best approximation of functional fitness would be a weighted sum of standardized scores. Therefore, categories should be created for already existing composite score rather than just be computed as a mean of scores already collapsed into categories as proposed by Buchman and his colleagues (Buchman, Boyle et al., 2007; Buchman, Wilson et al., 2007) and others (for details see the previous chapter). This novel approach ensures that less of original information is lost and that results are more accurate and most importantly, it is according to the theory. Because there will hardly be a case in which a person would not be able to perform any of SFT test, the five point scale was recommended (especially if the distribution is normal or close to normal) such it was in our case (figure 6.1).

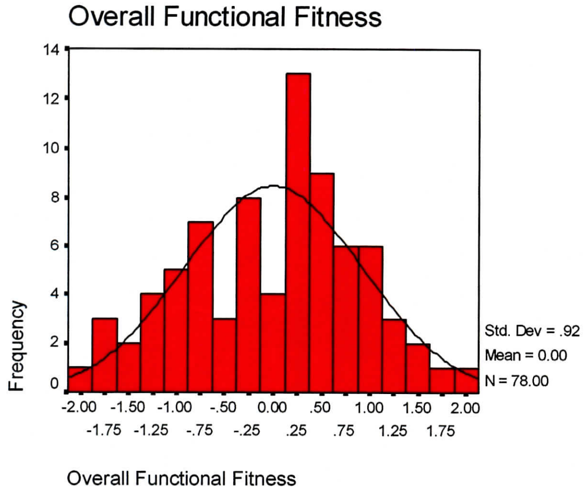


Figure 5.1 Distribution of Overall Functional Fitness scores (skewnes = -0.239, kurtosis = 0.539)

The very last table illustrates two real examples of all earlier proposed possible ways of data interpretation. Two completely different subjects were selected: the first one was a lower functioning participant and the second one was a higher functioning participant. Categories for individual tests were done using quartiles for the study sample hence they ranged from 1 to 4 where 4 represented the best result and 0 was added to represent those who were not able to perform. Categories for the Overall Functional Fitness were done using 20^{th}, 40^{th}, 60^{th}, and 80^{th} percentiles for the study sample so categories ranged from 1 to 5 where 1 represented the worst possible score and 5 exactly the opposite.

Table 5.1 Practical illustration of different ways of data interpretation for two extreme subjects

	Lower functioning subject		**Higher functioning subject**	
	Original units	**Categories (0-4 for individual tests and 1-5 for composite score**	**Original units**	**Categories (0-4 for individual tests and 1-5 for composite score**
Test 1	3 repetitions	1	11 repetitions	4
Test 2	4 repetitions	1	13 repetitions	3
Test 3	22 repetitions	3	54 repetitions	4
Test 4	-31 cm	1	4 cm	4
Test 5	-----	0	1 cm	4
Test 6	101 seconds	2	9.1 seconds	3
Overall Functional Fitness	-1.48 SD	1	1.14 SD	5

CHAPTER 6 - References and bibliography

Alexander, J. L., Phillips, W. T., & Wagner, C. L. (2008). The effect of strength training on functional fitness in older patients with chronic lung disease enrolled in pulmonary rehabilitation. *Rehabil Nurs, 33*(3), 91-97.

Avila-Funes, J. A., Gray-Donald, K., & Payette, H. (2006). [Measurement of physical capacities in the elderly: a secondary analysis of the Quebec longitudinal study NuAge]. *Salud Publica Mex, 48*(6), 446-454.

Baker, M. K., Atlantis, E., & Fiatarone Singh, M. A. (2007). Multi-modal exercise programs for older adults. *Age Ageing, 36*(4), 375-381.

Bearden, W. O., Sharma, S., & Teel, J. E. (1982). Sample size effects on chi-square and other statistics used in evaluating casual models. *Journal of marketing research, 19*, 425-430.

Beck, A. M., Damkjaer, K., & Beyer, N. (2008). Multifaceted nutritional intervention among nursing-home residents has a positive influence on nutrition and function. *Nutrition*.

Bentler, P. M. (1986). Structural modeling in psychometrika: an historical perspective on growth and achievements. *Psychometrica, 51*, 35-51.

Blyth, F. M., Cumming, R., Mitchell, P., & Wang, J. J. (2007). Pain and falls in older people. *Eur J Pain, 11*(5), 564-571.

Bollen, K. A. (1989). *Structural equations with latent variables*. New york: A Wiley-Interscience Publications.

Bollen, K. A., & Long, J. S. (1993). *Testing Structural equation models*. Newbury Park: Sage Publications.

Booth, F. W., Laye, M. J., Lees, S. J., Rector, R. S., & Thyfault, J. P. (2008). Reduced physical activity and risk of chronic disease: the biology behind the consequences. *Eur J Appl Physiol, 102*(4), 381-390.

Boyle, P. A., Buchman, A. S., Wilson, R. S., Bienias, J. L., & Bennett, D. A. (2007). Physical activity is associated with incident disability in community-based older persons. *J Am Geriatr Soc, 55*(2), 195-201.

Brach, J. S., Simonsick, E. M., Kritchevsky, S., Yaffe, K., & Newman, A. B. (2004). The association between physical function and lifestyle activity and exercise in the health, aging and body composition study. *J Am Geriatr Soc, 52*(4), 502-509.

Brach, J. S., VanSwearingen, J. M., Newman, A. B., & Kriska, A. M. (2002). Identifying early decline of physical function in community-dwelling older women: performance-based and self-report measures. *Phys Ther, 82*(4), 320-328.

Branch, L. G., Katz, S., Kniepmann, K., & Papsidero, J. A. (1984). A prospective study of functional status among community elders. *Am J Public Health, 74*(3), 266-268.

Brown, M., Sinacore, D. R., Ehsani, A. A., Binder, E. F., Holloszy, J. O., & Kohrt, W. M. (2000). Low-intensity exercise as a modifier of physical frailty in older adults. *Arch Phys Med Rehabil, 81*(7), 960-965.

Bruce, R. A., Kusumi, F., & Hosmer, D. (1973). Maximal oxygen intake and nomographic assessment of functional aerobic impairment in cardiovascular disease. *Am Heart J, 85*(4), 546-562.

Buchman, A. S., Boyle, P. A., Wilson, R. S., Bienias, J. L., & Bennett, D. A. (2007). Physical activity and motor decline in older persons. *Muscle Nerve, 35*(3), 354-362.

Buchman, A. S., Wilson, R. S., Boyle, P. A., Bienias, J. L., & Bennett, D. A. (2007). Change in motor function and risk of mortality in older persons. *J Am Geriatr Soc, 55*(1), 11-19.

Cao, Z. B., Maeda, A., Shima, N., Kurata, H., & Nishizono, H. (2007). The effect of a 12-week combined exercise intervention program on physical performance and gait kinematics in community-dwelling elderly women. *J Physiol Anthropol, 26*(3), 325-332.

Capodaglio, P., Capodaglio Edda, M., Facioli, M., & Saibene, F. (2007). Long-term strength training for community-dwelling people over 75: impact on muscle function, functional ability and life style. *Eur J Appl Physiol, 100*(5), 535-542.

Carr, K., Emes, C., & Rogerson, M. (2003). Exercise testing protocols for different abilities in the older population. *Activities, Adaptation, and Aging, 28*(1), 49-66.

Carvalho, M. J., Marques, E., & Mota, J. (2008). Training and Detraining Effects on Functional Fitness after a Multicomponent Training in Older Women. *Gerontology*.

Conn, V. S., Minor, M. A., Burks, K. J., Rantz, M. J., & Pomeroy, S. H. (2003). Integrative review of physical activity intervention research with aging adults. *J Am Geriatr Soc, 51*(8), 1159-1168.

Coste, J., Fermanian, J., & Venot, A. (1995). Methodological and statistical problems in the construction of composite measurement scales: a survey of six medical and epidemiological journals. *Stat Med, 14*(4), 331-345.

Coste, J., Walter, E., & Venot, A. (1995). A new approach to selection and weighting of items in evaluative composite measurement scales. *Stat Med, 14*(23), 2565-2580.

Cress, E., Ferrucci, L., Studenski, S., Guralnik, J. M., Fried, L. P., Gill, T. M., et al. (2008). Functional outcomes for clinical trials in frail older persons: time to be moving. *J Gerontol A Biol Sci Med Sci, 63*(2), 160-164.

Cress, M. E., Buchner, D. M., Questad, K. A., Esselman, P. C., deLateur, B. J., & Schwartz, R. S. (1996). Continuous-scale physical functional performance in healthy older adults: a validation study. *Arch Phys Med Rehabil, 77*(12), 1243-1250.

Cress, M. E., Petrella, J. K., Moore, T. L., & Schenkman, M. L. (2005). Continuous-scale physical functional performance test: validity, reliability, and sensitivity of data for the short version. *Phys Ther, 85*(4), 323-335.

Cyarto, E. V., Brown, W. J., Marshall, A. L., & Trost, S. G. (2008). Comparison of the effects of a home-based and group-based resistance training program on functional ability in older adults. *Am J Health Promot, 23*(1), 13-17.

Daley, M. J., & Spinks, W. L. (2000). Exercise, mobility and aging. *Sports Med, 29*(1), 1-12.

de Vreede, P. L., van Meeteren, N. L., Samson, M. M., Wittink, H. M., Duursma, S. A., & Verhaar, H. J. (2007). The effect of functional tasks exercise and resistance exercise on health-related quality of life and physical activity. A randomised controlled trial. *Gerontology, 53*(1), 12-20.

DiBrezzo, R., Shadden, B. B., Raybon, B. H., & Powers, M. (2005). Exercise intervention designed to improve strength and dynamic balance among community-dwelling older adults. *J Aging Phys Act, 13*(2), 198-209.

DiPietro, L. (2001). Physical activity in aging: changes in patterns and their relationship to health and function. *J Gerontol A Biol Sci Med Sci, 56 Spec No 2*, 13-22.

Dobek, J. C., White, K. N., & Gunter, K. B. (2006). The effect of novel ADL-based training program on performance of activities of daily living and physical fitness. *Journal of Aging and Physical Activity, 15*, 13-25.

Evans, W. J. (1995). *Keys to successful aging* Paper presented at the International Conference on Aging and Physical Activity

Ferrer, M., Lamarca, R., Orfila, F., & Alonso, J. (1999). Comparison of performance-based and self-rated functional capacity in Spanish elderly. *Am J Epidemiol, 149*(3), 228-235.

Ferrucci, L., & Simonsick, E. M. (2006). A little exercise. *J Gerontol A Biol Sci Med Sci, 61*(11), 1154-1156.

Fiantore Singh, M. A. (2002). Exercise comes of Age: Rationale and recommendations for a Geriatric Exercise Prescription. *Journal of Gerontology: Medical Sciences, 57A* (5), 262-282.

Fox, K. R. (1999). The influence of physical activity on mental well-being. *Public Health Nutr, 2*(3A), 411-418.

Frankel, J. E., Bean, J. F., & Frontera, W. R. (2006). Exercise in the elderly: research and clinical practice. *Clin Geriatr Med, 22*(2), 239-256; vii.

Fried, L. P., Bandeen-Roche, K., Chaves, P. H., & Johnson, B. A. (2000). Preclinical mobility disability predicts incident mobility disability in older women. *J Gerontol A Biol Sci Med Sci, 55*(1), M43-52.

Garatachea, N., Molinero, O., Martinez-Garcia, R., Jimenez-Jimenez, R., Gonzalez-Gallego, J., & Marquez, S. (2008). Feelings of well being in elderly people: Relationship to physical activity and physical function. *Arch Gerontol Geriatr*.

Gerbing, D. W., & Anderson, J. C. (1993). Monte Carlo evaluations of goodness-of-fit indices for structural equation models. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 320). Newbury Park: Sage Publications.

Gerety, M. B., Mulrow, C. D., Tuley, M. R., Hazuda, H. P., Lichtenstein, M. J., Bohannon, R., et al. (1993). Development and validation of a physical performance instrument for the functionally impaired elderly: the Physical Disability Index (PDI). *J Gerontol, 48*(2), M33-38.

Guralnik, J. M., Ferrucci, L., Pieper, C. F., Leveille, S. G., Markides, K. S., Ostir, G. V., et al. (2000). Lower extremity function and subsequent disability: consistency across studies, predictive models, and value of gait speed alone compared with the short physical performance battery. *J Gerontol A Biol Sci Med Sci, 55*(4), M221-231.

Guralnik, J. M., Ferrucci, L., Simonsick, E. M., Salive, M. E., & Wallace, R. B. (1995). Lower-extremity function in persons over the age of 70 years as a predictor of subsequent disability. *N Engl J Med, 332*(9), 556-561.

Guralnik, J. M., LaCroix, A. Z., Abbott, R. D., Berkman, L. F., Satterfield, S., Evans, D. A., et al. (1993). Maintaining mobility in late life. I. Demographic characteristics and chronic conditions. *Am J Epidemiol, 137*(8), 845-857.

Guralnik, J. M., Seeman, T. E., Tinetti, M. E., Nevitt, M. C., & Berkman, L. F. (1994). Validation and use of performance measures of functioning in a non-disabled older population: MacArthur studies of successful aging. *Aging (Milano), 6*(6), 410-419.

Guralnik, J. M., Simonsick, E. M., Ferrucci, L., Glynn, R. J., Berkman, L. F., Blazer, D. G., et al. (1994). A short physical performance battery assessing lower extremity function: association with self-reported disability and prediction of mortality and nursing home admission. *J Gerontol, 49*(2), M85-94.

Hall, K. M. (1997). The Functional Assessment Measure (FAM). *J Rehabil Outcomes, 1*(3), 63-65.

Haskell, W. L., Lee, I. M., Pate, R. R., Powell, K. E., Blair, S. N., Franklin, B. A., et al. (2007). Physical activity and public health: updated recommendation for adults from the

American College of Sports Medicine and the American Heart Association. *Med Sci Sports Exerc, 39*(8), 1423-1434.

Hauer, K., Becker, C., Lindemann, U., & Beyer, N. (2006). Effectiveness of physical training on motor performance and fall prevention in cognitively impaired older persons: a systematic review. *Am J Phys Med Rehabil, 85*(10), 847-857.

Hollmann, W., Struder, H. K., Tagarakis, C. V., & King, G. (2007). Physical activity and the elderly. *Eur J Cardiovasc Prev Rehabil, 14*(6), 730-739.

Holmerova, I. (2007). *Supporting Family Carers of Older People in Europe –The National Background Report for the Czech Republic*. Hamburg: LIT Verlag.

Holmerova, I., Juraskova, B., & Zikmundova, K. (2003). *Vybrane kapitoly z gerontologie* (2 ed.). Praha: Cas.

Joreskog, K. G. (1969). A general approach to confirmatory maximum likelihood factor analysis. *Psychometrica, 34*, 183-202.

Joreskog, K. G., & Sorbom, D. (1989). User's reference guide [Scientific Software]. Chicago, IL.

Judge, J. O., Lindsey, C., Underwood, M., & Winsemius, D. (1993). Balance improvements in older women: effects of exercise training. *Phys Ther, 73*(4), 254-262; discussion 263-255.

Jylha, M., Guralnik, J. M., Balfour, J., & Fried, L. P. (2001). Walking difficulty, walking speed, and age as predictors of self-rated health: the women's health and aging study. *J Gerontol A Biol Sci Med Sci, 56*(10), M609-617.

Kalvach, Z., Janeckova, H., & Bures, I. (2004). Long Term health and social care in institutions. In Z. Kalvach, Z. Zadak & R. Jirak (Eds.), *Geriatrics and Gerontology* (pp. 490-507). Prague: Grada Publishing.

Katz, S., Branch, L. G., Branson, M. H., Papsidero, J. A., Beck, J. C., & Greer, D. S. (1983). Active life expectancy. *N Engl J Med, 309*(20), 1218-1224.

Katz, S., Ford, A. B., Moskowitz, R. W., Jackson, B. A., & Jaffe, M. W. (1963). Studies of Illness in the Aged. The Index of Adl: A Standardized Measure of Biological and Psychosocial Function. *JAMA, 185*, 914-919.

Keith, R. A., Granger, C. V., Hamilton, B. B., & Sherwin, F. S. (1987). The functional independence measure: a new tool for rehabilitation. *Adv Clin Rehabil, 1*, 6-18.

Kinugasa, T., & Nagasaki, H. (1998). Reliability and validity of the Motor Fitness Scale for older adults in the community. *Aging (Milano), 10*(4), 295-302.

Kirk, K. L. (2001). Dietary restriction and aging: comparative tests of evolutionary hypotheses. *J Gerontol A Biol Sci Med Sci, 56*(3), B123-129.

Kivinen, P., Sulkava, R., Halonen, P., & Nissinen, A. (1998). Self-reported and performance-based functional status and associated factors among elderly men: the Finnish cohorts of the Seven Countries Study. *J Clin Epidemiol, 51*(12), 1243-1252.

Kramer, A. F., & Erickson, K. I. (2007). Capitalizing on cortical plasticity: influence of physical activity on cognition and brain function. *Trends Cogn Sci, 11*(8), 342-348.

Kruger, J., Bowles, H. R., Jones, D. A., Ainsworth, B. E., & Kohl, H. W., 3rd. (2007). Health-related quality of life, BMI and physical activity among US adults (>/=18 years): National Physical Activity and Weight Loss Survey, 2002. *Int J Obes (Lond), 31*(2), 321-327.

Lawton, M. P., & Brody, E. M. (1969). Assessment of older people: self-maintaining and instrumental activities of daily living. *Gerontologist, 9*(3), 179-186.

Lemmink, K. A. P. M., Han, K., de Greef, M. H. G., Rispens, P., & Stevens, M. (2001). Reliability of the Groningen Fitness Test for the elderly. *Journal of aging and physical activity 9*, 194-212.

Machacova, K., Bunc, V., Vankova, H., Holmerova, I., & Veleta, P. (2007). Zkusenosti s hodnocením telesne zdatnosti senioru metodou "Senior Fitness Test"

Ceska geriatricka revue, 4.

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

Mahrova, A., Bunc, V., & Fischerova, H. (2006). [Motor skills testing in patients with chronic renal failure]. *Cas Lek Cesk, 145*(10), 782-787.

Mobbs, C. V., Bray, G. A., Atkinson, R. L., Bartke, A., Finch, C. E., Maratos-Flier, E., et al. (2001). Neuroendocrine and pharmacological manipulations to assess how caloric restriction increases life span. *J Gerontol A Biol Sci Med Sci, 56 Spec No 1*, 34-44.

Morley, J. E. (2003). Hot topics in geriatrics. *J Gerontol A Biol Sci Med Sci, 58*(1), 30-36.

Myers, A. M., Holliday, P. J., Harvey, K. A., & Hutchinson, K. S. (1993). Functional performance measures: are they superior to self-assessments? *J Gerontol, 48*(5), M196-206.

Nagasaki, H., Itoh, H., & Furuna, T. (1995a). A physical fitness model of older adults. *Aging (Milano), 7*(5), 392-397.

Nagasaki, H., Itoh, H., & Furuna, T. (1995b). The structure underlying physical performance measures for older adults in the community. *Aging (Milano), 7*(6), 451-458.

Netz, J., & Argot, E. (1997). Assessment of functional fitness among independent older adults: a preliminary report. *Percept Mot Skills, 84*, 1059-1074.

Nieves, J. W., Zion, M., Pahor, M., Bernabei, R., Gussekloo, J., Simon, H., et al. (2005). Evaluation of continuous summary physical performance scores (CSPPS) in an elderly cohort. *Aging Clin Exp Res, 17*(3), 193-200.

Nocon, M., Hiemann, T., Muller-Riemenschneider, F., Thalau, F., Roll, S., & Willich, S. N. (2008). Association of physical activity with all-cause and cardiovascular mortality: a systematic review and meta-analysis. *Eur J Cardiovasc Prev Rehabil, 15*(3), 239-246.

Onder, G., Penninx, B. W., Lapuerta, P., Fried, L. P., Ostir, G. V., Guralnik, J. M., et al. (2002). Change in physical performance over time in older women: the Women's Health and Aging Study. *J Gerontol A Biol Sci Med Sci, 57*(5), M289-293.

Osness, W. H., Adrian, M., & Clark, B. (1996). *Functional fitness assessment for adults over 60 years*. Dubuque: Kendal: Hunt Publishing.

Ostir, G. V., Volpato, S., Fried, L. P., Chaves, P., & Guralnik, J. M. (2002). Reliability and sensitivity to change assessed for a summary measure of lower body function: results from the Women's Health and Aging Study. *J Clin Epidemiol, 55*(9), 916-921.

Paterson, D. H., Jones, G. R., & Rice, C. L. (2007). Ageing and physical activity: evidence to develop exercise recommendations for older adults. *Can J Public Health, 98 Suppl 2*, S69-108.

Poehlman, E. T., Turturro, A., Bodkin, N., Cefalu, W., Heymsfield, S., Holloszy, J., et al. (2001). Caloric restriction mimetics: physical activity and body composition changes. *J Gerontol A Biol Sci Med Sci, 56 Spec No 1*, 45-54.

Reuben, D. B., Laliberte, L., Hiris, J., & Mor, V. (1990). A hierarchical exercise scale to measure function at the Advanced Activities of Daily Living (AADL) level. *J Am Geriatr Soc, 38*(8), 855-861.

Reuben, D. B., & Siu, A. L. (1990). An objective measure of physical function of elderly outpatients. The Physical Performance Test. *J Am Geriatr Soc, 38*(10), 1105-1112.

Reuben, D. B., Valle, L. A., Hays, R. D., & Siu, A. L. (1995). Measuring physical function in community-dwelling older persons: a comparison of self-administered, interviewer-administered, and performance-based measures. *J Am Geriatr Soc, 43*(1), 17-23.

Rikli, R. E., & Jones, J. C. (2001). *Senior fitness test manual*. Champain, IL: Human Kinetics.

Sawatzky, R., Liu-Ambrose, T., Miller, W. C., & Marra, C. A. (2007). Physical activity as a mediator of the impact of chronic conditions on quality of life in older adults. *Health Qual Life Outcomes, 5*, 68.

Segars, A. H. (1997). Assessing the unidimensionality of measurement: a paradigm and illustration within the context of information systems research. *Int J Mgmt Sci, 25*(1), 107-121.

Shepard, R., J (Ed.). (1997). *Aging, physical activity, and health*. Champain, IL: Human Kinetics.

Sherman, S. E., & Reuben, D. (1998). Measures of functional status in community-dwelling elders. *J Gen Intern Med, 13*(12), 817-823.

Shibata, A., Oka, K., Nakamura, Y., & Muraoka, I. (2007). Recommended level of physical activity and health-related quality of life among Japanese adults. *Health Qual Life Outcomes, 5*, 64.

Spirduso, W. W. (Ed.). (2005). *Physical dimension of aging* (2nd ed ed.). Champain, IL: Human Kinetics.

Stewart, K. J. (2005). Physical activity and aging. *Ann N Y Acad Sci, 1055*, 193-206.

Ware, J. E., Jr., & Sherbourne, C. D. (1992). The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care, 30*(6), 473-483.

Wilson, T. M., & Tanaka, H. (2000). Meta-analysis of the age-associated decline in maximal aerobic capacity in men: relation to training status. *Am J Physiol Heart Circ Physiol, 278*(3), H829-834.

Winograd, C. H., Lemsky, C. M., Nevitt, M. C., Nordstrom, T. M., Stewart, A. L., Miller, C. J., et al. (1994). Development of a physical performance and mobility examination. *J Am Geriatr Soc, 42*(7), 743-749.

Wood, R. H., Gardner, R. E., Ferachi, K. A., King, C., Ermolao, A., Cherry, K. E., et al. (2005). Physical function and quality of life in older adults: sex differences. *South Med J, 98*(5), 504-512.

Appendix

A1 Lisrel output

DATE: 1/29/2009
TIME: 16:02

L I S R E L 8.50

BY

Karl G. Jöreskog & Dag Sörbom

This program is published exclusively by
Scientific Software International, Inc.
7383 N. Lincoln Avenue, Suite 100
Lincolnwood, IL 60712, U.S.A.
Phone: (800)247-6113, (847)675-0720, Fax: (847)675-2140

Website: www.ssicentral.com

The following lines were read from file C:\Users\machacova\Documents\Lisrel fitness model.spj:

Title
Senior Fitness Test

Observed Variables
test1 test2 test3 test4 test5 test6

Correlation Matrix

1
0.543 1
0.595 0.426 1
0.533 0.431 0.454 1
0.356 0.337 0.399 0.399 1
-0.511 -0.325 -0.514 -0.481 -0.458 1

Sample size
78

Latent Variable
Fitness

Paths
Fitness -> test1 test2 test3 test4 test5 test6

Lisrel Output ME = ML RS PC SC MI WP FS AD = off IT = 999

Path Diagram
End of Problem

Senior Fitness Test

Correlation Matrix

	test1	test2	test3	test4	test5	test6
test1	1.00					
test2	0.54	1.00				
test3	0.59	0.43	1.00			
test4	0.53	0.43	0.45	1.00		
test5	0.36	0.34	0.40	0.40	1.00	
test6	-0.51	-0.33	-0.51	-0.48	-0.46	1.00

Senior Fitness Test

Parameter Specifications

LAMBDA-X

	Fitness
test1	1
test2	2
test3	3
test4	4
test5	5
test6	6

THETA-DELTA

test1	test2	test3	test4	test5	test6
7	8	9	10	11	12

Senior Fitness Test

Number of Iterations = 5

LISREL Estimates (Maximum Likelihood)

LAMBDA-X

Fitness

----

test1	0.79
	(0.10)
	7.69
test2	0.61
	(0.11)
	5.53
test3	0.73
	(0.11)
	6.89
test4	0.68
	(0.11)
	6.27
test5	0.54
	(0.11)
	4.78
test6	-0.67
	(0.11)
	-6.23

PHI

Fitness
1.00

THETA-DELTA

test1	test2	test3	test4	test5	test6
0.38	0.62	0.47	0.54	0.70	0.54
(0.09)	(0.11)	(0.10)	(0.10)	(0.12)	(0.10)
4.28	5.51	4.89	5.22	5.72	5.24

Squared Multiple Correlations for X - Variables

test1	test2	test3	test4	test5	test6
0.62	0.38	0.53	0.46	0.30	0.46

Goodness of Fit Statistics

Degrees of Freedom = 9

Minimum Fit Function Chi-Square = 7.64 (P = 0.57)
Normal Theory Weighted Least Squares Chi-Square = 7.64 (P = 0.57)
Estimated Non-centrality Parameter (NCP) = 0.0
90 Percent Confidence Interval for NCP = (0.0 ; 9.04)

Minimum Fit Function Value = 0.099
Population Discrepancy Function Value (F0) = 0.0
90 Percent Confidence Interval for F0 = (0.0 ; 0.12)
Root Mean Square Error of Approximation (RMSEA) = 0.0
90 Percent Confidence Interval for RMSEA = (0.0 ; 0.11)
P-Value for Test of Close Fit (RMSEA < 0.05) = 0.70

Expected Cross-Validation Index (ECVI) = 0.43
90 Percent Confidence Interval for ECVI = (0.43 ; 0.55)
ECVI for Saturated Model = 0.55
ECVI for Independence Model = 2.16

Chi-Square for Independence Model with 15 Degrees of Freedom = 154.69
Independence AIC = 166.69
Model AIC = 31.64
Saturated AIC = 42.00
Independence CAIC = 186.83
Model CAIC = 71.93
Saturated CAIC = 112.49

Normed Fit Index (NFI) = 0.95
Non-Normed Fit Index (NNFI) = 1.02
Parsimony Normed Fit Index (PNFI) = 0.57
Comparative Fit Index (CFI) = 1.00
Incremental Fit Index (IFI) = 1.01
Relative Fit Index (RFI) = 0.92

Critical N (CN) = 219.39

Root Mean Square Residual (RMR) = 0.038
Standardized RMR = 0.038
Goodness of Fit Index (GFI) = 0.97
Adjusted Goodness of Fit Index (AGFI) = 0.93
Parsimony Goodness of Fit Index (PGFI) = 0.41

Senior Fitness Test

Fitted Covariance Matrix

	test1	test2	test3	test4	test5	test6
test1	1.00					
test2	0.48	1.00				
test3	0.58	0.45	1.00			
test4	0.54	0.42	0.49	1.00		
test5	0.43	0.33	0.40	0.37	1.00	

test6	-0.53	-0.41	-0.49	-0.46	-0.37	1.00

Fitted Residuals

	test1	test2	test3	test4	test5	test6
test1	0.00					
test2	0.06	0.00				
test3	0.02	-0.02	0.00			
test4	0.00	0.01	-0.04	0.00		
test5	-0.07	0.00	0.00	0.03	0.00	
test6	0.02	0.09	-0.02	-0.02	-0.09	0.00

Summary Statistics for Fitted Residuals

Smallest Fitted Residual = -0.09
Median Fitted Residual = 0.00
Largest Fitted Residual = 0.09

Stemleaf Plot

```
- 0|97
- 0|4222000000000
  0|1223
  0|69
```

Standardized Residuals

	test1	test2	test3	test4	test5	test6
test1	- -					
test2	1.48	- -				
test3	0.68	-0.43	- -			
test4	-0.07	0.27	-0.93	- -		
test5	-1.68	0.04	0.04	0.49	- -	
test6	0.64	1.62	-0.52	-0.48	-1.49	- -

Summary Statistics for Standardized Residuals

Smallest Standardized Residual = -1.68
Median Standardized Residual = 0.00
Largest Standardized Residual = 1.62

Stemleaf Plot

```
- 1|75
- 0|9554100000000
  0|3567
  1|56
```

Senior Fitness Test

Qplot of Standardized Residuals

No Non-Zero Modification Indices for LAMBDA-X

No Non-Zero Modification Indices for PHI

Modification Indices for THETA-DELTA

	test1	test2	test3	test4	test5	test6
test1	- -					
test2	2.19	- -				
test3	0.46	0.19	- -			
test4	0.01	0.07	0.87	- -		
test5	2.83	0.00	0.00	0.24	- -	
test6	0.41	2.61	0.27	0.23	2.22	- -

Expected Change for THETA-DELTA

	test1	test2	test3	test4	test5	test6
test1	- -					
test2	0.11	- -				
test3	0.05	-0.03	- -			
test4	-0.01	0.02	-0.07	- -		
test5	-0.13	0.00	0.00	0.04	- -	
test6	0.05	0.13	-0.04	-0.04	-0.12	- -

Completely Standardized Expected Change for THETA-DELTA

	test1	test2	test3	test4	test5	test6
test1	- -					
test2	0.11	- -				
test3	0.05	-0.03	- -			
test4	-0.01	0.02	-0.07	- -		
test5	-0.13	0.00	0.00	0.04	- -	
test6	0.05	0.13	-0.04	-0.04	-0.12	- -

Maximum Modification Index is 2.83 for Element (5, 1) of THETA-DELTA

Covariance Matrix of Parameter Estimates

	LX 1,1	LX 2,1	LX 3,1	LX 4,1	LX 5,1	LX 6,1	TD 1,1	TD 2,2	TD 3,3	TD 4,4
LX 1,1	0.01									
LX 2,1	0.00	0.01								
LX 3,1	0.00	0.00	0.01							
LX 4,1	0.00	0.00	0.00	0.01						
LX 5,1	0.00	0.00	0.00	0.00	0.01					
LX 6,1	0.00	0.00	0.00	0.00	0.00	0.01				
TD 1,1	0.00	0.00	0.00	0.00	0.00	0.00	0.01			
TD 2,2	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01		
TD 3,3	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	
TD 4,4	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01

TD 5,5	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
TD 6,6	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Covariance Matrix of Parameter Estimates

	TD 5,5	TD 6,6
TD 5,5	0.02	
TD 6,6	0.00	0.01

Senior Fitness Test

Correlation Matrix of Parameter Estimates

	LX 1,1	LX 2,1	LX 3,1	LX 4,1	LX 5,1	LX 6,1	TD 1,1	TD 2,2	TD 3,3	TD 4,4
LX 1,1	1.00									
LX 2,1	0.25	1.00								
LX 3,1	0.31	0.23	1.00							
LX 4,1	0.29	0.21	0.26	1.00						
LX 5,1	0.22	0.17	0.20	0.19	1.00					
LX 6,1	-0.28	-0.21	-0.26	-0.24	-0.19	1.00				
TD 1,1	-0.28	0.04	0.06	0.05	0.03	-0.05	1.00			
TD 2,2	0.03	-0.17	0.02	0.01	0.01	-0.01	-0.05	1.00		
TD 3,3	0.06	0.02	-0.24	0.03	0.02	-0.03	-0.10	-0.02	1.00	
TD 4,4	0.04	0.01	0.02	-0.21	0.01	-0.02	-0.07	-0.02	-0.04	1.00
TD 5,5	0.02	0.01	0.01	0.01	-0.15	-0.01	-0.03	-0.01	-0.02	-0.01
TD 6,6	0.04	0.01	0.02	0.02	0.01	0.20	-0.07	-0.02	-0.04	-0.02

Correlation Matrix of Parameter Estimates

	TD 5,5	TD 6,6
TD 5,5	1.00	
TD 6,6	-0.01	1.00

Senior Fitness Test

Factor Scores Regressions

KSI

	test1	test2	test3	test4	test5	test6
Fitness	0.32	0.15	0.24	0.19	0.12	-0.19

Senior Fitness Test

Standardized Solution

LAMBDA-X

Fitness

	Fitness
test1	0.79
test2	0.61
test3	0.73
test4	0.68
test5	0.54
test6	-0.67

PHI

Fitness
1.00

Senior Fitness Test

Completely Standardized Solution

LAMBDA-X

	Fitness
test1	0.79
test2	0.61
test3	0.73
test4	0.68
test5	0.54
test6	-0.67

PHI

Fitness
1.00

THETA-DELTA

test1	test2	test3	test4	test5	test6
0.38	0.62	0.47	0.54	0.70	0.54

Time used: 0.031 Seconds