Application of Methods in Physical Activity Measurement

by

Stephen Herrmann

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved April 2011 by the
Graduate Supervisory Committee:

Barbara Ainsworth, Chair
Glenn Gaesser
Cheryl Der Ananian
Minsoo Kang
Sonia Vega-Lopez

ARIZONA STATE UNIVERSITY

May 2011

ABSTRACT

It is broadly accepted that physical activity provides substantial health benefits. Despite strong evidence, approximately 60% to 95% of US adults are insufficiently active to obtain these health benefits. This dissertation explored five projects that examined the measurement properties and methodology for a variety of physical activity assessment methods. Project one identified validity evidence for the new MyWellness Key accelerometer in sixteen adults. The MyWellness Key demonstrated acceptable validity evidence when compared to a criterion accelerometer during graded treadmill walking and in free-living settings. This supports the use of the MyWellness Key accelerometer to measure physical activity. Project two evaluated validity (study 1) and test-retest reliability evidence (study 2) of the Global Physical Activity Questionnaire (GPAQ) in a two part study. The GPAQ was compared to direct and indirect criterion measures including object and subjective physical activity instruments. These data provided preliminary validity and reliability evidence for the GPAQ that support its use to assess physical activity. Project three investigated the optimal $h \cdot d^{-1}$ of accelerometer wear time needed to assess daily physical activity. Using a semi-simulation approach, data from 124 participants were used to compare 10-13 $h \cdot d^{-1}$ to the criterion 14 $h \cdot d^{-1}$. This study suggested that a minimum accelerometer wear time of 13 $h \cdot d^{-1}$ is needed to provide a valid measure of daily physical activity. Project four evaluated validity and reliability evidence of a novel method

(Movement and Activity in Physical Space [MAPS] score) that combines accelerometer and GPS data to assess person-environment interactions. Seventy-five healthy adults wore an accelerometer and GPS receiver for three days to provide MAPS scores. This study provided evidence for use of a MAPS score for future research and clinical use. Project five used accelerometer data from 1000 participants from the 2005-2006 National Health and Nutrition Examination Study. A semi-simulation approach was used to assess the effect of accelerometer wear time (10-14 $\text{h}\cdot\text{d}^{-1}$) on physical activity data. These data showed wearing for 12 $\text{h}\cdot\text{d}^{-1}$ or less may underestimate time spent in various intensities of physical activity.

DEDICATION

This dissertation is dedicated to my family. My wife, Sarah, for the encouragement and sacrifice she made during my time in this graduate program. In addition to supporting us financially during my years of full-time study, she shared equally in all of the emotional and financial burdens involved. Next is my son, Brody, who was born near the beginning of this process. His innocence and happiness provided numerous moments that allowed me to forget about the countless projects, papers, and stressors that go along with doctoral work. Lastly this is dedicated, to my mom and dad, for instilling the importance of hard work and higher education.

ACKNOWLEDGENT

I wish to thank my committee members who were more than generous with their expertise and valuable time. A special thanks to Dr. Barbara Ainsworth, my committee chair, for her countless hours of reflecting, reading, encouraging, and most of all patience throughout the entire process. Thank you to Dr. Glenn Gaesser, Dr. Cheryl Der Ananian, Dr. Minsoo Kang, and Dr. Sonia Vega-Lopez for agreeing to serve on my committee.

I would also like to acknowledge and thank other members of Exercise and Wellness and Nutrition for guiding me through the years and providing any assistance requested. A final thanks goes to Julie Rice for her friendship and immeasurable help.

TABLE OF CONTENTS

CHAPTER Page

LIST OF TABLES

Table Page

Table Page

LIST OF FIGURES

Figure Page

Chapter 1

INTRODUCTION

It is widely understood that physical activity provides substantial health benefits. For years studies have shown that men, women, and ethnically diverse individuals, who are moderately active on a regular basis, have lower mortality rates than those who are inactive (Haskell et al., 2007; Mokdad, Marks, Stroup, & Gerberding, 2004; Paffenbarger, Hyde, Wing, & Hsieh, 1986; Pate et al., 1995; Wannamethee, Shaper, & Walker, 1998; Warburton, Nicol, & Bredin, 2006). Recently, studies have provided more detail to support the role of physical activity in a variety of health conditions and the overall benefit in public health. The research clearly shows that physical activity could serve as the primary preventive behavior for several major health problems such as hypertension (Padilla, Wallace, & Park, 2005), dyslipidemia (Kraus et al., 2002; Murphy, Nevill, Neville, Biddle, & Hardman, 2002), diabetes (Gregg, Gerzoff, Caspersen, Williamson, & Narayan, 2003; LaMonte, Blair, & Church, 2005), and cardiovascular disease (Li et al., 2006; Murphy et al., 2002).

To improve understanding of the health benefits of physical activity, it is important to understand the definition of physical activity. Physical activity is a behavioral construct that is often confused with exercise, energy expenditure, and/or physical fitness (Caspersen, Powell, & Christenson, 1985; LaPorte et al., 1984). A commonly used definition for physical activity includes “any bodily movement produced by skeletal muscles that results in energy expenditure” (Caspersen et al., 1985). Physical activity is typically characterized by mode or type of activity (e.g., walking, running, sweeping), frequency (how often), duration (minutes), and intensity (light, moderate, vigorous). The energy necessary to conduct these physical activities is described and measured as the energy cost of physical activity or energy expenditure (LaMonte & Ainsworth, 2001). When a person participates in a structured or systematic program, it is referred to as exercise. Appropriate amounts of exercise lead to maintenance or improvement in cardiovascular fitness. There is substantial literature to support the importance of cardiovascular fitness as it relates to health and mortality (Blair et al., 1996; Carnethon, Gulati & Greenland, 2005; Laukkanen et al., 2001; Lee, Blair, & Jackson, 1999; Ruiz et al., 2007)

Regular physical activity has many proven health benefits which has led to several organizations providing physical activity recommendations or guidelines. Guidelines are statements about the type, intensity, frequency and duration of physical activity intended to help people improve fitness, maintain or improve health, and reduce the risk of a variety of chronic diseases. In 1995, the Centers for Disease Control and Prevention and the American College of Sports Medicine provided a joint statement on physical activity and health (Pate, Pratt, et al., 1995). This statement provided concise information on the type and intensity of activity for health promotion and disease prevention. The recommendation stated that every adult should accumulate 30 minutes or more of moderate intensity physical activity on most, but preferably all, days of the week (Pate, Pratt, et al., 1995). Moderate physical activity was defined as activities performed at an intensity of approximately 3 to 6 METs (Metabolic Equivalent) or comparable to that of brisk walking at 3 to 4 mph (Ainsworth, Haskell, et al., 1993).

More recently, the United States Department of Health and Human Services provided the first formal US government physical activity guidelines (USDHHS, 2008). In lieu of daily dosage, the report titled, “Physical Activity Guidelines for Americans” presented the recommended dose as an accumulation moderate and/or vigorous activity across the entire week. Specifically, adults should participate in at least 150 minutes per week of moderate intensity activity, 75 minutes of vigorous intensity activity, or comparable combination with more activity (300 minutes of moderate, 150 minutes of vigorous, or combined equivalent) providing an added health benefit (“Physical Activity Guidelines Advisory Committee Report.,” 2008). Furthermore, the report offered additional guidelines for older adults, youth, and other special groups (e.g., persons with disabilities, during pregnancy, etc).

Physical inactivity poses a serious public health problem. The global prevalence of individuals that fall below recommended levels of physical activity is reported near 60% (“WHO | Physical activity,” n d). The

World Health Organization (WHO) estimates that worldwide 2 million deaths annually are attributed to physical inactivity related chronic diseases such as cancer, diabetes, and heart disease (“WHO | Physical activity,” n d). In the United States, estimates of adults meeting physical activity guidelines range from a low of less than 5% to as high as 45% (Macera, Ham, et al., 2005; Troiano et al., 2007). By engaging in regular physical activity many adults can improve or maintain their health and reduce their risks of chronic disease and premature mortality (Haskell, Lee, et al., 2007). Low levels of regular physical activity leaves many individuals at greater risk for physical inactivity related chronic diseases such as hypertension, dyslipidemia, insulin resistance (impaired glucose tolerance), cancers, coronary heart disease, stroke, type 2 diabetes, osteoporosis and osteoarthritis (Popkin, Rusev, Du, & Zizza, 2006; Warburton, Nicol, & Bredin, 2006).

Large differences in the prevalence of physical activity in populations may be attributed to differences in measurement methods used to assess physical activity. Objective methods used to assess physical activity include motion detectors (e.g., accelerometers, pedometers) and subjective measures (e.g., questionnaires, records, and logs) (Welk, 2002; Wood & Zhu, 2006). Objective measures reflect actual movement and are often preferred to quantify total movement by volume and/or intensity and duration. New accelerometers are being developed and marketed for use in research, clinical, and practice settings, but oftentimes little is known about their validity and reliability evidence. The ActiGraph accelerometer is used most frequently as an objective measure of physical activity, however some questions still exist, including understanding the number of hours per day the monitor should be worn to accurately reflect time spent in varying intensities of movement.

Subjective instruments provide a recall of physical activity types and doses, are relatively inexpensive, and predominantly are used for population physical activity surveillance and in epidemiological study designs (Blair, Haskell, Ho, et al., 1985; Macera, Ham, et al., 2005; Paffenbarger, Wing, & Hyde, 1978). Questionnaires have been developed to identify national and global prevalence of physical activity. However, some of these questionnaires have not been sufficiently evaluated in terms of their validity and reliability evidence. Evaluation of physical activity measurement instruments is important to assure adequate validity and reliability for research studies designed to assess physical activity levels in populations and to identify changes in physical activity behaviors.

Use of the Global Positioning System (GPS) has recently become more feasible for assessing physical activity. Since May 2000, when the U.S. government turned off selective availability (i.e., an intentional amount of error in the GPS), the accuracy of GPS has improved significantly (White House Office of the Press Secretary, 2000). Recent work using GPS units have validated the accuracy of GPS to assess various speeds of walking, running, and cycling (Le Faucheur et al., 2007;

Schutz & Herren, 2000). This has led to others using GPS to assess walking ability in disease populations and to measure active transport of healthy individuals in an urban environment (Duncan, Mummery, & Dascombe, 2007; Le Faucheur et al., 2008). However, because GPS is a new area, standard protocols do not yet exist.

An emerging area in physical activity assessment is the integration of multiple physical activity measures (e.g., heart rate monitors and accelerometer – Actiheart, etc.)(Brage, Brage, Franks, Ekelund, & Wareham, 2005; Tapia et al., 2007). The Actiheart, for example, uses heart rate and activity data from an accelerometer to predict activity energy expenditure with similar accuracy to that of indirect calorimetry (Crouter, Churilla, & Bassett, 2008). Integrating Global Positioning Systems (GPS) and accelerometers may help to understand where individuals engage in physical activity (Herrmann et al., 2008; Quigg, Gray, Reeder, Holt, & Waters, 2010; Troped, Oliveira, Matthews, et al., 2008; Troped, Wilson, Matthews, Cromley, & Melly, 2010). Physical activity space conceptualizes this and is defined as the area, or space, within which an individual spends time and engages in physical activities (Zhu, 2003). Little is known about how determine their physical activity space or if there is an association between physical activity space and health outcomes.

This dissertation is a compilation of five separate research projects with the overall theme of physical activity measurement and evaluation.

The studies were designed to identify the validity and reliability of objective and subjective physical activity measures, examine differences in the accelerometer wear time to assess the effect on physical activity data, and to describe the descriptive epidemiology of physical activity space.

Statement of the Problem

An enduring problem in the assessment of physical activity is the different scores provided when using subjective and/or objective tools to measure physical activity participation. These scores can alter the prevalence of physical activity levels in surveillance settings and cause doubt in the effectiveness of research studies to provide the desired effect. For example, U.S. physical activity surveillance systems traditionally used physical activity questionnaires to assess population activity levels. In the 2001 Behavioral Risk Factors Surveillance System (BRFSS), using a telephone questionnaire, 40% of adults were classified as sufficiently active to meet public health recommendations (Macera, Ham, et al., 2005); whereas in the 2003-2004 National Health and Nutrition Examination Study (NHANES) that used accelerometers to assess time spent in physical activity sufficient to meet public health recommendations, only 5% of adults were deemed active (Troiano et al., 2007). While subjective instruments are often inexpensive to administer and provide additional context to an individual's physical activity, they have a high risk for recall bias rendering them less accurate (Welk, 2002). In order for questionnaires to be accurate, they must be valid and reliable. The Global Physical Activity Questionnaire (GPAQ) is used by the World Health Organization to assess the global prevalence of health-enhancing physical activity (Armstrong & Fiona Bull, 2006). This survey has very limited validity and reliability evidence (Armstrong & Fiona Bull, 2006) and has not been examined for accuracy against accelerometers and other questionnaires used in global physical activity surveillance. Alternatively, objective measures of physical activity, such as accelerometers, provide recordings of time spent in activity at various intensities, but do not provide any context of the activity and are generally limited to measuring only ambulatory movement (Welk, 2002). Due to the popularity of accelerometry based measures to assess physical activity, companies are developing and marketing accelerometers for used by the public to monitor exercise routines and daily physical activity. However, little is known about the validity and reliability evidence of these new accelerometers. Additionally, a persistent question regarding accelerometer use is how long must they be worn per day to reflect time spent in physical activity of varying intensities? There is little evidence to support current recommendations for the amount of time one should wear an accelerometer to reflect a typical day's physical activity. It is imperative that the methodology and instruments employed have sound scientific support with detailed information regarding their measurement properties. This dissertation reflects a series of studies that address the accuracy of physical activity measurement using objective and subjective measures and explore the assessment of physical activity space.

Hypotheses

Five projects addressed the properties of instruments and methodologies used to assess physical activity behavior. Research hypotheses for each project are listed below.

Project One. Evaluation of the MyWellness Key Accelerometer.

purpose. To evaluate the validity of the Technogym MyWellness accelerometer.

Project One Hypotheses

1A. The association between the MyWellness accelerometer and a criterion accelerometer is similar at three physical activity intensity levels (light, moderate, vigorous) on a treadmill in a laboratory setting.

1B. The MyWellness accelerometer provides similar physical activity data compared to other objective and subjective measures under free-living conditions

Project Two. Validity of the Global Physical Activity Questionnaire.

purpose. To assess the validity and reliability of the Global Physical Activity Questionnaire (GPAQ).

Project Two Hypotheses

2A. There is an association between the GPAQ, objective measures of physical activity, and a subjective measure of physical activity.

2B. There is categorical agreement between the GPAQ and the IPAQ when classifying individuals by physical activity level.

2C. The GPAQ is a reliable measure of inactivity, and moderate and vigorous physical activity for each domain (travel, recreation, work) of the questionnaire.

Project Three. How Many Hours are Enough? Optimal Accelerometer Wear Time to Reflect Daily Physical Activity

purpose. To determine the optimal ActiGraph accelerometer wear time needed to reflect daily physical activity.

Project Three Hypotheses

3A. Increased accelerometer daily wear time produces more accurate assessments of daily physical activity.

3B. There is significantly less physical activity demonstrated when fewer less wear time is allowed.

Project Four. Validity of the Movement and Activity in Physical Space (MAPS) score in Healthy Adults.

purpose. To examine the validity and reliability characteristics of MAPS scores measured by combining global positioning systems and accelerometry in healthy adults.

Project Four Hypotheses

4A. Higher Movement and Activity in Physical Space (MAPS) values are associated with higher levels of physical activity (measured by accelerometer) and environmental interaction (measured by GPS).

4B. Three days of monitoring with MAPS provides a reliable estimate of free-living function.

4C. MAPS scores demonstrate convergent and discriminant validity with other activity, neighborhood, and health related constructs.

Project Five. Impact of Accelerometer Wear Time on Physical Activity Data.

purpose. To evaluate the effects of varying amounts of accelerometer wear time on physical activity in the NHANES study.

Project Five Hypotheses

5A. Time spent in light, moderate, and vigorous intensity physical activity differs depending on the wear time criteria.

5B. Sedentary time increases with increased daily accelerometer wear time.

5C. Increased daily accelerometer wear time produces more accurate assessments of daily physical activity compared to the criterion wear time.

Scope

The projects included in this dissertation are designed to identify information about the measurement and evaluation of physical activity and the application of a variety of methods. These projects described the measurement properties of both subjective and objective tools used to assess physical activity, evaluated a novel method for evaluating physical activity space, and investigated methodologies for improving accelerometer measurement techniques. The samples for studies 1 and 4 were generally healthy adults with wide ranging physical activity levels. Studies 2 and 3 included university employees. Study 5 included a larger probability sample to reflect the United States population.

Objective physical activity instruments included two accelerometers and a pedometer used in laboratory and free-living conditions. Subjective physical activity instruments included two short recall questionnaires, and a 72-hour physical activity record. Additional subjective instruments were used to assess neighborhood health, social support and self efficacy for exercise, general health, and a travel log. Measures of free-living function included the use of a GPS unit and geographic information system (GIS) software. The physical fitness measures utilized were estimated maximal oxygen uptake (VO_2 max) and steady state heart rate obtained from a submaximal cycle ergometer protocol, resting hear rate, and resting systolic blood pressure and diastolic blood pressure. Anthropometric measures consisted of percent fat, body mass index, waist circumference, and sagittal diameter.

Assumptions

1. Three to seven days of physical activity monitoring by accelerometers is an accurate measure of physical activity behavior.

2. Participants recorded all of their physical activity on physical activity questionnaires and in physical activity records to the best of their ability.
3. Participants accurately followed instructions for placement and wear of physical activity monitoring devices.
4. Participants were truthful in completing study surveys and completed all study protocols to the best of their ability.

Limitations

1. Waist circumference, percent fat, BMI, estimated VO_2 max, steady state heart rate, and resting blood pressure were measured accurately and reflect indirect criterion measures for physical activity.
2. The quasi-experimental cross-sectional design for projects 1, 4, and 5 preclude casual inferences.

Significance of the Research

Despite the strong evidence which supports physical activity as a primary preventive measure to decrease risk in a variety of health outcomes, approximately 60% to 95% of US adults are insufficiently active to obtain these health benefits (Haskell et al., 2007; Macera et al., 2005; Troiano et al., 2007). Since reports of physical activity vary greatly, it is important to improve knowledge of existing, new, and emerging tools to attain a greater understanding of their measurement capabilities along with developing new methodologies to improve assessment techniques. This dissertation 1) examined the measurement properties of objective and subjective physical activity instruments; 2) investigated a novel approach for understanding free-living function and; 3) explored methodology in accelerometer assessment to improve estimates of physical activity.

Definitions

1. <u>Cardiorespiratory fitness:</u> A health related component of physical fitness that relates to the ability of the circulatory and respiratory systems to supply oxygen during sustained physical activity (US Department of Health and Human Services, 1996).

2. <u>1995 CDC-ACSM Physical Activity Recommendation:</u> Adults should obtain ≥ 30 minutes per day of moderate physical activity (3-6 METs) on most, if not all, days of the week (Pate et al., 1995).

3. <u>1996 Surgeon General's Report – Physical Activity Recommendation:</u> Adults should accumulate at least 150 kcal per day or 1,000 kcal per week of moderate (3-6 METs) and/or vigorous physical activity (> 6 METs) (USDHHS, 1996).

4. 2007 ACSM-AHA Physical Activity Recommendation: Adults aged 18 to 65 years need moderate intensity aerobic physical activity for a minimum of 30 minutes on five days each week or vigorous intensity aerobic physical activity for a minimum of 20 minutes on three days each week. Combinations of moderate- and vigorous intensity activity can be performed to meet this recommendation (Haskell et al., 2007).

5. 2008 USDHHS Physical Activity Recommendation (Adults): Adults should do at least 150 minutes a week of moderate intensity activity, 75 minutes a week of vigorous intensity activity, or an equivalent combination. Aerobic activity should be performed in episodes of at least 10 minutes, and preferably, it should be spread throughout the week. For additional health benefits, adults should increase their aerobic physical activity to 300 minutes a week of moderate intensity activity, 150 minutes a week of vigorous intensity activity or an equivalent combination. Adults should also do muscle-strengthening activities that are moderate or high intensity and involve all major muscle groups on 2 or more days a week ("Physical Activity Guidelines Advisory Committee Report.," 2008).

6. Metabolic equivalent (MET): A MET is a unit used to estimate the metabolic cost (oxygen consumption) of physical activity. For an average adult; one MET equals the resting metabolic rate (sitting quietly), which is approximately 3.5 ml of oxygen per kg body weight per minute or 1 kcal per kg body weight per hour (Ainsworth et al., 1993).

7. Physical Activity: Any bodily movement produced by skeletal muscles that results in energy expenditure (Caspersen et al., 1985).

8. Physical Activity Space: The area, or space, within which an individual spends time and engages in physical activities (Zhu, 2003).

Chapter 2

LITERATURE REVIEW

The association between physical activity and health is not new. In the ninth century BC writing, *Regimen in Health*, Hippocrotes stated that “exercise should be many and of all kinds….quiet to begin with, increase till they are violent and then gently finishing (Lagerros & Lagiou, 2007). This review of literature examines a variety of topics related to physical activity recommendations and the measurement of physical activity.

Physical Activity Recommendation

Regular physical activity has many proven health benefits which has led to several organizations providing physical activity recommendations or guidelines. These statements, intended for adults, offer information to help people improve fitness, maintain or improve health, and reduce the risk of a variety of chronic diseases.

Early position statements and guidelines for physical activity generally focused on exercise and cardiovascular fitness. In 1975, the American College of Sports Medicine published “Guidelines for Graded Exercise Testing and Exercise Prescription” (ACSM, 1975) with five later revisions updating each edition with new evidence (ACSM, 1980; ACSM, 1986; ACSM, 1991; ACSM, 1995).

In the 1990's physical activity became a top public health issue (Pratt, Epping, & Dietz, 2009). In 1995, the Centers for Disease Control and Prevention and the American College of Sports Medicine provided a joint statement on physical activity and health (Pate et al., 1995). The purpose of this document was to present a concise "public health message" with information on the type and intensity of activity for health promotion and disease prevention (Pate et al., 1995). The message by Pate et al. (1995) was that every adult should accumulate 30 minutes or more of moderate intensity physical activity on most, but preferably all, days of the week. One vital component of this recommendation was the importance of moderate intensity physical activity to achieve health benefits. Moderate physical activity is activity performed at an intensity of approximately 3 to 6 METs (Metabolic Equivalent) or comparable to that of brisk walking at 3 to 4 mph (Ainsworth et al., 1993). A second important aspect of this statement was that physical activity accumulated in short intermittent bouts is a suitable method to achieve these activity goals and obtain the associated health benefits (Pate et al., 1995). These recommendations have been adopted widely throughout many other countries (Oja, Bull, Fogelholm, & Martin, 2010).

In 1996, the Surgeon General's Report on Physical Activity and Health followed up with similar recommendations (USDHHS, 1996). This report emphasized that health benefits can be obtained from a "moderate" level of activity which was defined as activity that uses 150 kcal per day or

1,000 kcal per week (USDHHS, 1996). The report acknowledges that vigorous levels of activity are not required for health benefits, but it can provide another method for achieving health benefits. Furthermore, it was concluded that regular physical can reduce the risk for developing a variety of disease (e.g., coronary heart disease, diabetes, hypertension, etc.) and helps maintain healthy bones, muscles, and joints (USDHHS, 1996).

The Institute of Medicine Committee (IOM) on Dietary Reference Intakes released a report on physical activity in 2002. The IOM advised 60 minutes of moderate intensity activity per day and described previous recommendations of 30 minutes per day to be inadequate. The IOM acknowledged that some benefits are possibly from 30 minutes of moderate intensity activity; however, this report has been criticized for its misrepresentation of current research data (Blair, LaMonte, & Nichaman, 2004)

By 2007, the American College of Sports Medicine and the American Heart Association provided an update to the 1995 physical activity recommendation for adults (Haskell, Lee, et al., 2007). The purpose of this update was to provide an update of the evidence and clarify certain aspects of the previous recommendation. The update recommended that healthy adults achieve at least 30 minutes of moderate intensity aerobic activity on five days a week or vigorous intensity aerobic activity for at least 20 minutes three days a week. It also stated that a combination of moderate and vigorous intensity activity can be used to meet these recommendations. Moreover, the recommendation states that additional health benefits may be achieved by exceeding these minimum recommendations. In addition, adults should perform muscular strength training activities at least twice a week. Of note, this updated report places greater focus on the benefit of vigorous intensity activity and activities for muscle and bone health than previous recommendations.

More recently, the United States federal government issued the first ever Physical Activity Guidelines for Americans (“Physical Activity Guidelines Advisory Committee Report.,” 2008). This report included recommendations for children through older adults and other special populations (e.g., people with disabilities, pregnant women, etc.). The basic guidelines are similar to the 2007 AHA-ACSM recommendations but focus on total weekly activity. For example, the guidelines specify at least 150 minutes per week of moderate intensity activity should be performed instead of 30 minutes, five days per week. This weekly recommendation is due in part to findings such as Lee et al. (2004) who identified men from the Harvard Alumni Health Study as “weekend warriors”. These men (without major risk factors) acquired all of their activity on 1-2 days while maintaining a lower risk of dying compared to sedentary men (95% confidence interval: 0.62, 0.91) (Lee, Sesso, Oguma, & Paffenbarger, 2004). The guidelines also state that seventy-five minutes of vigorous activity can be done to meet the guidelines, or a combination of moderate and vigorous activity. Additional health benefits are possibly for those who double the recommendations (300 minutes of moderate or 150 minutes of vigorous, or an equivalent combination). The broad theme of these guidelines is that some activity is better than no activity at all, health benefits increase with increased dose (i.e., intensity, duration, frequency), and the benefits of physical activity overshadows the risks.

Assessing Physical Activity

Physical activity is most commonly measured by subjective self-reports and objective activity monitoring instruments.

Subjective Measures

Self-reported physical activity instruments include activity logbooks, diaries, recall questionnaires, and voice recorders. These self-report options utilize a person’s ability to recall their activities from a specified time period (e.g., 24 hours, 7 days, 1 year, etc). Self-report instruments are relatively inexpensive, generally easy to administer and complete, and oftentimes provide additional information about the context of the physical activity. Specific subjective instruments generally fall under a few broad types, including recall, global, quantitative history, and logbooks.

Recall questionnaires. Recall questionnaires are the most common form of physical activity assessment in epidemiologic studies (LaMonte &

Ainsworth, 2001). In general, recall questionnaires provide information about physical activity over the past 24 hours, week, or month. In about 10-20 items, recall questionnaires supply data about the type, frequency, duration, and intensity of activity (Craig et al., 2003). This information is useful in physical activity surveillance and behavior studies to identify individuals that are meeting activity guidelines, should be targeted for intervention, or are successful in improving or changing behavior.

A recently developed and commonly used recall questionnaire is the International Physical Activity Questionnaire (IPAQ). The IPAQ was developed to evaluate physical activity and sedentary behaviors using a variety of physical activity domains with a short format (time in sedentary, moderate, and vigorous intensity activity and time spent walking) and a long format (time in leisure, work, household, yard, and sedentary activity, as well as self-powered transport) that could be used by telephone interview or self-administered (Bauman et al., 2009; Craig et al., 2003; Hallal et al., 2010; Macfarlane, Lee, Ho, Chan, & Chan, 2007). The IPAQ has been translated into approximately 20 languages and has been used in many countries worldwide (Brown, Trost, Bauman, Mummery, & Owen, 2004; Craig et al., 2003). Reliability studies have shown wide ranging values for test-retest reliability of the IPAQ (0.34 to 0.93) (Craig, Marshall, et al., 2003; Ekelund et al., 2006; Hagströmer, Oja, & Sjöström, 2006). Multiple validation studies have been published on the IPAQ yielding generally positive yet relatively modest correlations with accelerometers (r

= 0.07 - 0.71) (Fogelholm et al., 2006; Rzewnicki, Vanden Auweele, & De Bourdeaudhuij, 2003).

Some researchers have questioned the accuracy of the IPAQ and have demonstrated substantial overreporting (Montoye, Kemper, Saris, & Washburn, 1996; Welk, 2002). For example, Rzewnicki et al. (2003) obtained physical activity information with the IPAQ and then followed up with probing participant interviews to better understand how participants responded. From these interviews, Rzewnicki found that nearly half of the sample reported some physical activity (walking, moderate-, or vigorous intensity) on the IPAQ when they should have reported no physical activity. Therefore, approximately 50% of individuals classified as meeting activity guidelines in fact, did not meet those guidelines because they over reported their activity. Furthermore, about 5% of respondents provided physical activity values so high that they were deemed not credible or impossible.

This over reporting of physical activity on physical activity questionnaires is not limited to the IPAQ. Self-reported instruments in general are often thought to suffer from inaccurate participant recall with associated errors ranging from 35% to 50% of recalled activities (Lagerros & Lagiou, 2007). Over reporting of physical activity is an issue to consider when using recall questionnaires. However, other types of physical activity self-report measures have their own limitations, such as activity records and activity logs, which oftentimes are thought to influence the physical activity patterns being measured (LaMonte & Ainsworth, 2001; Matthews, Ainsworth, Hanby, et al., 2005).

Global questionnaires. Global questionnaires are characterized by being rather short; having one to four items (Godin & Shephard, 1985; LaMonte & Ainsworth, 2001; Shephard & Vuillemin, 2003). While this form of subjective physical activity assessment lessens the participant burden it also provides only limited information about a person's physical activity. Due to the brevity, global questionnaires typically lack details regarding patterns of physical activity and specific time in different activity intensity levels. The primary purpose of global questionnaires is to provide only broad classifications (e.g., active or inactive) (Matthews, 2002).

Quantitative History. Quantitative history questionnaires require a more comprehensive format to acquire detailed information. These questionnaires typically evaluate physical activity using 15 to 60 items to assess frequency, intensity, and duration of activity during the past month, year, or lifetime (Jacobs, Ainsworth, Hartman, & Leon, 1993; Slattery & Jacobs, 1995; Taylor, Jacobs, et al., 1978). The Minnesota Leisure Time Physical Activity Questionnaire was one of the first questionnaires to provide an extensive list of activities (63 sports, recreational, yard, and household activities) (Lagerros & Lagiou, 2007). The purpose of this interviewer administered questionnaire format is to obtain the type and frequency of activities to calculate an estimate of energy expenditure. The time and cost of training interviewers, contacting participants, and coding the data are significant drawbacks of this form of in-depth interviewer assisted recall (Ainsworth, Haskell, et al., 1993; Bouchard et al., 1983).

Records and logbooks. Physical activity logbooks, records and/or diaries are used to obtain detail on the type and duration of all the activities someone performs. Using records or diaries, the participant is instructed to keep a record of all the activities as they occur or at a specified time interval (e.g., 15 minutes) (Ainsworth, Haskell, et al., 1993). This generated list of activities can be coded according to energy expenditure or MET value to identified physical activity patterns and understand behavior (LaMonte, Ainsworth, & Reis, 2006; Washburn, Heath, & Jackson, 2000). This format is burdensome to the participant and requires considerable effort by the researcher to code entries. Despite these limitations, physical activity records are often considered a criterion measure for subjective physical activity assessment (LaMonte et al., 2006).

Logbooks, on the other hand, provide a checklist of activities performed in the past day or during the day. This reduces some participant and researcher burden but may not be inclusive of all physical activities performed during the day, thus, missing activities that are actually performed (Lagerros & Lagiou, 2007; LaMonte & Ainsworth, 2001). Logbooks and records can potentially alter participants’ physical activity behaviors due to their increased awareness (Centers for Disease Control and Prevention (CDC). National Center for Health Statistics (NCHS), 2010).

Objective Measures

Objective measures of physical activity include pedometers, accelerometers, heart rate monitors, indirect calorimetry and doubly labeled water. These instruments and methods objectively quantify activity or physiological responses to physical activity. When used appropriately, objective measures are helpful in quantifying physical activity; however, there are some limitations in their abilities to record all aspects of physical activity, such as context or type and location of movement. The cost of these objective instruments can vary ranging from only a couple of dollars to thousands of dollars. Some instruments require expert knowledge and special software to use and evaluate the data. Additionally, these devices can increase the burden on an individual by requiring them to wear a monitor attached to their body or clothing.

Activity monitors. Pedometers and accelerometers have become an increasingly popular objective method for physical activity assessment. A recent search in PubMed returned 1,696 articles for “accelerometer” and 603 articles for “pedometer” published from January 2001 to January

2011. The search during 1990 through 2000 returned only 75 articles for “pedometer” and 378 articles for “accelerometer”.

Pedometers. Pedometers have become a widely used measure for assessing walking (Bassett, Wyatt, Thompson, Peters, & Hill, 2010; Beets, Bornstein, Beighle, Cardinal, & Morgan, 2010; Craig, Cameron, Griffiths, & Tudor-Locke, 2010). These devices often range in price from only a few dollars up to $200 depending on the internal mechanism and device options (i.e., memory, software, etc) (Schneider, Crouter, & Bassett, 2004). The primary outcome measure is a step count which allows for some pedometers to estimate distance walked (from stride length [distance = steps × stride length]), energy expended in movement, and activity intensity (from step rate in $\text{steps}\cdot\text{min}^{-1}$). Pedometers are generally accurate in counting steps within 3% of actual steps taken (Schneider, Crouter, Lukajic, & Bassett, 2003) and become less accurate for estimating distance, and even less accurate for estimating energy expenditure (Crouter, Schneider, Karabulut, & Bassett, 2003b). However, significant variation can be found due to the internal mechanism and sensitivity causing under- or overestimation by 25% - 45% (Schneider, Crouter, & Bassett, 2004; Schneider, Crouter, Lukajic, et al., 2003). To minimize this error, it is important for researchers and clinicians to be aware of the reliability and validity evidence for the devices they use.

Pedometers function from a few different internal mechanisms; horizontal spring-lever, magnetic reed proximity switch, or a piezo-electric mechanism. The horizontal spring-lever mechanism responds to vertical movements at the hip by swinging the lever up and down to close an electrical circuit (Schneider, Crouter, et al., 2003). When the lever makes contact, it counts the steps and oftentimes produces an audible “click” which can be an easy way to distinguish this type of pedometer mechanism.

The magnetic reed proximity switch is comprised of a magnet connected to a spring suspended horizontal lever arm which responds to vertical hip movement. Steps are counted when a magnetic field is created that activates a proximity switch inside of a glass cylinder (Schneider, Crouter, Lukajic, & Bassett, 2003). These two mechanisms make use of coiled or hairspring mechanisms that have the potential to wear out over time affecting sensitivity.

The third common mechanism used in pedometers is a piezo-electric mechanism. The piezo-electric mechanism uses a strain gauge to measure inertia that interprets step count and is better than the previously described mechanisms at estimating activity intensity and energy expenditure. This type of device also improves measurement accuracy when walking at slower speeds that might fail to register steps with a spring levered mechanism (Crouter, Schneider, Karabulut, & Bassett, 2003).

A common public health message using pedometers is to encourage adults to achieve 10,000 steps per day (Hatano, 1993). Walking for 10,000 step per day expends about 300 kcal of added energy expenditure, reflects the dose of physical activity identified as optimal for reducing the risk of having a first heart attack (Paffenbarger, Kampert, & Lee, 1997), and is generally associated with a healthy level of physical activity (Chan, Ryan, & Tudor-Locke, 2004; Wilde, Sidman, & Corbin, 2001). Step counts goals have also been compared to physical activity recommendations that encourage 30 minutes per day of moderate intensity activity. This recommendations can be achieved by accumulating 3,000 – 4,000 steps that are of moderate intensity (≥ 100 steps per minute) (Tudor-Locke, Sisson, Collova, Lee, & Swan, 2005), accumulated in at least 10 minute bouts, and are above a sedentary level of physical activity (e.g., 5,000 steps per day) (Tudor-Locke, Hatano, Pangrazi, & Kang, 2008). In 2004, a “zone-based hierarchy” was identified for pedometer step count indices (Tudor-Locke & Bassett, 2004). Table 1 displays the preliminary classifications for daily walking activity in healthy adults.

Table 1

Activity Classification Based on Steps per Day

Classification	Steps/Day
Sedentary	< 5,000
Low Active	5,000 - 7,499
Somewhat Active	7,500 - 9,999
Active	10,000 – 12,499
Highly Active	> 12,500

Accelerometers. Accelerometers are used to assess the body's motion (i.e., acceleration) as a result of movement and physical activity. The majority of accelerometers fall into two categories (Uniaxial and Triaxial) based on their ability to assess activity in a single or multiple movement planes. Uniaxial accelerometers measure accelerations in a single (i.e., vertical) axis. Triaxial accelerometers measure body accelerations in three planes of movement (i.e., vertical, horizontal, and lateral) planes. The accelerations are interpreted as intensity of movement.

Accelerometers are comparable in size to a pedometer and are a very useful tool when conduction free-living and field physical activity research. One advantage of accelerometers is that they are constantly sampling movement and lack of movement providing detailed (e.g., minute-by-minute) output on activity intensity. From this constant output, accelerometers offer information on the frequency and duration of movement by intensity levels.

Most accelerometers are designed to accurately measure ambulatory activity, thus, when worn on the waist, accelerometers may underestimate activities such as weight lifting, bicycling, isometric exercise, and other activities that produce less body movement but expend energy. The high cost of some devices may limit the use of accelerometers in large scale surveillance studies. However, the National Health and Nutrition Examination Survey began using accelerometers in 2003 as part of its ongoing study to assess the health and nutritional status of adults and children in the United States (Haggett, 1965).

Measurement Concerns in Physical Activity

Obtaining accurate and reliable measures of physical activity is the foundation for research and practice in the field of physical activity. In physical activity assessment, errors can arise from a variety of sources; therefore, it is of great importance to minimize measurement error by using the best methods available. For that reason, the instruments used and their scores must have certain qualities that provide evidence for their use (i.e., evidence of validity and reliability).

Validity

The following is not an exhaustive list of type of validity but provides an overview of commonly used methods in physical activity assessment. Validity is often referred to as the most important or most fundamental concept for developing and evaluating instruments or scores (American Psychological Association, 1999, p. 9). Obtaining validity evidence is thought to be an ongoing process rather than a yes or no quality of an instrument or a score (Sheppard, 1993). Establishing validity evidence is context specific and therefore should be focused on the intended use and interpretation of test scores before an instrument is recommended (Rowe and Maher, 2006).

Understanding the underlying construct validity, or construct being measured, (e.g., physical activity) is central to understanding validity. This is done by considering different types of validity evidence. Criterion-related validity evidence is identified by comparing scores from the instrument of interest with a criterion measure. In the area of physical activity, it is difficult to identify a true "gold standard" measure (Sallis and Saelens, 2000; Shepherd and Vuillemin, 2003). Concurrent evidence of validity is most frequently used in physical activity research with accelerometers and pedometers along with measures of cardiorespiratory fitness and anthropometry. These surrogate measures have strong construct validity evidence to support their use as a surrogate to evaluate physical activity. This is because physical activity, as a construct, is thought to have many non-overlapping dimensions (Jacobs et al, 1993). For example, Craig et al. (2003) investigated criterion validity evidence of the IPAQ by comparing the results of 781 participants with physical activity data from accelerometers and found a correlation of $r = .30$. Modest associations are common in physical activity measures which may be due to the instruments providing imprecise estimates, or the validation measures do not account for all dimensions of physical activity (Papathanasiou et al., 2010).

Convergent validity evidence is also used in physical activity by using different measures to assess the same construct (physical activity) and evaluate the relationship of these different measures. In a review of 25 articles on convergent validity evidence for pedometers, Tudor-Locke et al. (2002) found results comparing pedometers to accelerometers, direct observation, and physical activity self-report. Pedometers were most strongly related to accelerometers (median r = 0.86) and direct observation (median r = 0.82) and the lowest correlation with self-report measures of physical activity (median r = 0.33) (Tudor-Locke, Williams, Reis, & Pluto, 2002). From this convergent validity evidence, this review suggested that the evidence supports pedometers as a valid alternative for physical activity assessment.

Known difference validity evidence is another common form of validity used in the field of physical activity. The known group difference method (Cronbach & Meehl, 1955) is used to assess the differences (e.g., physical activity) between groups (e.g., normal versus obese) or evaluate changes in physical activity pre- and post intervention to assess meaningful change in physical activity. For example, Morinder et al. (2009) evaluated known group difference validity evidence for the six minute walk test in obese and normal weight children (Morinder, Mattsson, Sollander, Marcus, & Larsson, 2009). These researchers found that the six minute walk test was able to satisfactorily differentiate between obese children (571 m) and normal weight (663 m) children ($p < 0.001$) (Morinder et al., 2009).

Reliability

While evaluating validity is considered to be exceptionally important, reliability is a “prerequisite” to studies of validity (Rowe & Maher, 2006). As a fundamental concept reliability is often described as consistency of measurement (i.e., same or similar score with multiple measures). Four types of reliability are discussed: Interclass reliability, intraclass reliability, test-retest reliability, and equivalence reliability.

Interclass reliability measures a bivariate relationship between variables and is based on the Pearson Product Moment correlation. Interclass reliability is assessed by comparing the results of the same, or similar, measure on two separate administrations. For example, the IPAQ has been evaluated using interclass reliability by administering the questionnaire on separate occasions (Jia, Xu, Kang, & Tang, 2008).

Intraclass reliability is similar to interclass reliability coefficient but allows for more than two trials/values to be analyzed. Intraclass reliability is founded within the analysis of variance (ANOVA) and is used to calculate the pooled reliability for all trials/values combined. This type of analysis has been used to identify the number of days an accelerometer should be worn to assess a person’s habitual activity. Matthews et al. (2002) used intraclass reliability to indentify the fewest number of days needed to obtain a reliable measure of physical activity using an accelerometer. They determined the reliability of 1, 3, 7, and 14 days of physical activity monitoring and found that a minimum of three to four days was needed for a reliable estimate of physical activity. Seven days of monitoring was needed to assess patterns of physical inactivity (Matthews, Ainsworth, Thompson, & Bassett, 2002).

Both interclass and intraclass reliability can be used to evaluate test-retest or equivalence reliability (Morrow, 2002). Test-retest reliability measures the correlation between separate administrations of the same test. For example, if a physical activity questionnaire is administered twice, two weeks apart, the correlation between the scores would be a test-retest reliability coefficient. Equivalence reliability is assessed when two similar instruments are used to assess the same construct. For example, if a new accelerometer was developed to assess physical activity, it could be compared with a current accelerometer that also assesses physical activity. If the two measures are correlated above a certain threshold (e.g., $r \geq 0.08$), then they are considered equivalent in their assessment of the same construct (i.e., physical activity).

How Many Days are Enough to Measure Habitual Physical Activity?

Reliability studies are common in physical activity research and are used to evaluate many different aspects including the reliability of instruments, scores, and behavior patterns. One area of interest is investigating the amount of monitoring needed to measure habitual physical activity.

Several studies have examined the number of days of physical activity monitoring needed to obtain a reliable estimate of habitual physical activity (Kang, Bassett, et al., 2009; Matthews, Hebert, et al., 2001; Matthews et al., 2002; Trost, Pate, Freedson, Sallis, & Taylor, 2000; Tudor-Locke, Burkett, Reis, et al., 2005). The purpose of these studies has been to identify the minimum number of days needed to measure physical activity to minimize participant burden. This is usually done using the intraclass correlation coefficient (ICC) with values greater than 0.08 being acceptable. The results of these studies are highly dependent on the variability of the data.

With a goal to determine the number of days needed to reflect daily step patterns, Tudor-Locke et al. (2005) performed a pedometer study in 90 adults with seven days of activity monitoring as the criterion. The results indicated that any three day combination of pedometer monitoring produced an ICC of 0.86 to 0.90. In another pedometer study with the goal of identifying how many days of monitoring is needed to reflect a year’s step pattern, Kang et al. (2009) had 23 adults wear a pedometer for 365 consecutive days as the criterion. This study demonstrated that at least five consecutive days or six random days were needed to obtain an ICC of greater than 0.80. Additionally, the results showed that at least 14 random days or 30 consecutive days were needed to obtain a mean absolute percentage error below the desired 10% (Kang et al., 2009).

Matthews et al. (2001) studied the sources of variance in self-reported physical activity. Using 24 hour physical activity recalls, Matthews et al. (2001) found that 50-60% of the variance in physical activity was due to within subject variation. Fifteen 24 hour recalls were given to 580 participants over 12 months. The results indicated that assessments of 7 to 10 days of 24-hour recalls in men and 14 to 21 days of 24-hour recalls in women were required to achieve a minimum reliability of 80% (Matthews et al., 2001).

Studies designed to identify the fewest days of accelerometer monitoring needed to assess usual physical activity patterns yield similar results as observed with pedometers. In children and adolescents, Trost et al. (2000) showed that children required four to five days of monitoring while adolescents needed eight to nine days to achieve appropriate levels of reliability to assess usual physical activity (Trost, Pate, Freedson, Sallis, & Taylor, 2000). In a study of adults, three days of accelerometer assessment was needed to reach a reliability of 0.80 ($r > 0.80$ is often considered desirable) for activity counts, four days of assessment for moderate to vigorous activity, and seven days for inactivity time

(Matthews, Freedson et al., 2001). A number of studies have investigated the necessary amount of days of physical activity monitoring needed to assess habitual activity. However, there has been little examination of the amount of time needed to identify a valid day of monitoring.

How Many Hours are Enough to Measure Daily Physical Activity?

The amount of daily accelerometer or pedometer wear time needed to obtain a valid estimate of daily activity is a similar issue to identifying the number of days needed to estimate habitual physical activity. However, there is little consensus in this area with researchers and clinicians using a variety of criteria.

Recent studies have reported a wide range of criteria for selecting a valid day (i.e., an accurate estimate of a single day's physical activity) from as few as 6 hours per day (Young et al., 2009) and up to 16 hours per day (Slootmaker et al., 2009) to constitute a valid day. Slootmaker et al. (2009) used an assumption that people sleep 8 hours per day and therefore restricted valid days to persons with less than 16 hours of monitoring. Results from a 2004 accelerometer consensus meeting, held in Chapel Hill North Carolina, suggested using the 70/80 rule (Catellier et al., 2005) for required daily wear time (Ward et al., 2005). This rule provides a sample specific recommendation based on 70% of the sample having accelerometer data. A valid day is then defined as 80% of that observed period. Another approach that has been used to determine accelerometer hours per day wear time is to normalize each person's total minutes of daily activity to 12 hours per day to balance different amounts of wear time (Young et al., 2009). For example, if a person has 15 minutes of moderate intensity activity over 10 hours of wear time, then there data would be adjusted to 18 minutes of moderate intensity activity over 12 hours. There are little data to support this method and the amount of error associated with this normalization is unclear as it may over- or under-estimate actual movement time. A common approach, including that used in U.S. National Health and Nutrition Examination Survey (NHANES) accelerometer analyses, is to require 10 or more hours per day of accelerometer wear time to be considered a valid measurement day (Matthews et al., 2008; Troiano et al., 2007). Improving accelerometer and pedometer methodology may help our understanding of physical activity behavior but these devices do not currently provide information about the context of activity. The concept of activity space may help us learn more about objectively measuring the context of activity.

Activity Space

The premise of physical activity space has its origins in the 1960's and 1970's. At this time researchers began to develop spatial concepts that have helped formulate our thinking today. In 1965, Haggett pioneered the idea of assessing locational geography and movement using nodes (locations), networks (travel paths) and areas, which is the foundation of much current research (Anderson, 1971; Haggett, 1965). Researchers in the 1970’s and 1980’s refined these ideas and focused on individual activity patterns and the value of space-time budgets (Palm, 1981), the influence of environment (Hagerstrand, 1970; Lenntorp, 1976), and the concept of Space-Time Paths and Prisms (Golledge & Stimson, 1997) which map or predict potential travel. For example, this can be used to understand how a person got to a location and if a person was in that location for 30 minutes, what were possible locations and likelihoods of them going different places during that allotted time?

Physical Activity Space

Physical activity space is derived from “activity space” proposed by Golledge and Stimson (Golledge & Stimson, 1997) which is described as the area where a person spends time. Golledge also used this term to describe “spatial behavior” which is another term that is often used interchangeable with activity space today. Weimo Zhu is a more recent proponent of activity space. In 2003, Zhu further refined the idea of Activity Space by calling it “Physical Activity Space” and defining it as “the area or space where an individual spends time *and engages in physical activity*” (Zhu, 2003). This definition of physical activity space includes three components; *time*, *space*, and *activity*. Nearly all current physical activity measures only measure activity, therefore, Zhu (2003) called for new measures of physical activity space that assess the interaction between an individual and the environment. Around the same time, Schönfelder (2002) was researching travel behavior and thought that spatial data analysis could help measure the concept of activity space and help identify how and why the environment influences behaviors (Schönfelder & Axhausen, 2002). Until recently, the built environment has been the focal point for the majority of health research using geospatial technology and its role in promoting physical activity. In 2008, Saarloos et al. proposed a bottom-up approach using activity-based modeling to understand how individuals interact in space and time with their environment and each other (Saarloos, Kim, & Timmermans, 2009). This approach may be very helpful in shifting the paradigm from a built environment approach to better understand individual spatial behavior.

Subjective methods, such as time-budget diaries (Anderson, 1971), have been used to assess Zhu's concept of physical activity space because direct measures have not been readily available. Of the three components of physical activity space (i.e., activity, space, and time), activity and time can be measure by activity logbooks or diaries by recording the time for every activity performed. An assessment of space could be appraised using a travel log. However, these types of methodologies are often limited by recall and classification difficulties.

Objective methods, such as pedometers and accelerometers, can be used in part to assess physical activity space. Recent advances in geospatial technology can assist in assessing space and time objectively using GPS receivers which automatically record time at a location and detailed travel data. Geographic Information Systems (GIS) can then be used to apply spatial statistics to analyze the data gathered from motion sensors and GPS. In theory, these technologies should be able to quantify physical activity space yet few people have proposed quantitative methodologies to do so (Herrmann & Ragan, 2008). When GPS is combined with accelerometers, it is possible to identify the frequency, duration, and intensity of an individual’s movement patterns within space.

Travel Behavior

The 1990’s was a popular time for studies investigating human travel behavior (Schönfelder & Axhausen, 2002; Wolf, Guensler, et al., 2001). Much of the work during the 1990’s was spent investigating differences in interpersonal and intrapersonal travel behavior. This field of study offers many parallels to physical activity behavior studies and can serve as a model for the physical activity field in the use of geospatial technology and spatial data analysis to identify where people engage in physical activity.

Schönfelder (2002) has investigated social exclusion based on travel and spatial behavior by exploring differences in groups that are at a higher risk for social exclusion (i.e., female, low SES, and elderly). Schönfelder proposed that these higher risk individuals take fewer trips and/or have less trip variety. However, the results of these hypotheses were inconclusive due to limited data on trip purpose and because the study was not designed to investigate such a relationship. Also of significant note, Schönfelder stated that this study lacked the ability to assess activity and energy expenditure of non-car travel which might help explain this relationship regarding travel and activity behaviors in individuals at higher risk for social exclusion (Schönfelder & Axhausen, 2002). Thus, additional research is needed to study mobility using spatial technologies in combination with physical activity measures.

Another example from travel studies comes from the Ratt Fart Project, a study conducted in Sweden from 1999 – 2001, that utilized an interactive approach to influence driving behaviors (Schoenfelder, Axhausen, Antille, & Bierlaire, 2002). Devices were placed in vehicles to evaluate speed, acceleration, and other travel data and then provide feedback to the driver about his or her performance in real-time. Combining this idea of real-time feedback with GPS/GIS and activity monitor technology could provide an interesting tool to influence physical activity and sedentary behaviors. For example, the device could know *not* to tell a person to get up and be active while they are commuting in a car or bus. When appropriate, this type of technology could provide real-time feedback regarding helpful tips/information and/or suggest nearby locations to be physically active such as parks, sidewalks, and fitness centers.

Combining Technology to Assess Physical Activity

Efforts are underway to combine GPS and accelerometer technology to improve the assessment of physical activity space (Patrick et al., 2008; Rodríguez et al., 2005). One of the earliest concept devices was designed in 1996 (Makikawa & Murakami, 1996). More recently, Rodriquez et al. (2005) investigated the capabilities of combining a GPS receiver and accelerometer to assess physical activity behavior but provided limited conclusions. The study reported that the device was able to measure activity and could measure location. However, there has not been any further development during the past few years reported in the literature about this device. Several researchers funded through the National Institute of Health Genes and Environment initiative are developing a system that uses multiple accelerometers and Bluetooth technology to send data to a smart phone (that is GPS capable) to interpret activity type (Patrick et al., 2008). The use of GPS capabilities has been limited in this system due to battery life and running multiple programs on the cell phone platform. However, this method offers a vast amount of potential for assessing physical activity behaviors and interaction with the environment.

Free-living Function

The Movement and Activity in Physical Space (MAPS) score is a newly developed method to quantify and provide and index score for physical activity space by combining data from separate GPS and accelerometer devices (Herrmann, Ragan, et al., 2008). This method incorporates the spatial assessment of the GPS by measuring locations where activities occur and matching by time with an accelerometer that measures characteristics about the activity (i.e., intensity, duration, etc.). Initial work using MAPS scores has demonstrated evidence of reliability and the ability to assess known-group differences in individuals with a reduced functional capacity and evidence of responsiveness to monitor their recovery/improvements in function over time (Herrmann et al., in press). Other researchers have adopted using MAPS scores to investigate other populations including individuals with multiple sclerosis (Snook, Scott, Ragan, Morand, & Sackau, 2010).

Limitations and Challenges in Physical Activity Space

While technology is rapidly advancing and multiple researchers are working on incorporating accelerometers with GPS and smart phone technology to enhance physical activity research, limitations and challenges exist. One challenge that exists is that physical activity researchers lack training in geospatial technology. Therefore, assessing spatial behavior may be best served through interdisciplinary approaches that include individuals from travel, geography, engineering, and physical activity. Another limitation of objectively measured activity space is that trip motives and/or purpose cannot always be easily assumed from GPS technology alone (Wolf, Guensler, et al., 2001). This provides an opportunity for researchers to include real-time querying of users (possibly with smart phones) to learn about activity (or trip) motives and/or purpose. In addition, Zhu (2003) identified that new spatial statistics need to be developed to better assess physical activity space. Lastly, practical problems also exist which include the relatively high cost of equipment and costly post-processing of GPS data to assess purpose and location.

Summary

Currently a variety of methods exist for physical activity assessment. However, it is essential that these instruments and methodologies be improved upon to advance understanding of physical activity assessment and behaviors. Doing so may help improve physical activity levels by developing and implementing more specific and targeted interventions.

Chapter 3

METHODS

This chapter provides an overview of the methods used for the five research projects presented as separate papers in Chapters 4-8. The research studies focused on a global theme of physical activity measurement and included the validation of a commercial accelerometer, evaluation of a physical activity questionnaire, description of physical activity and environment interaction, and identification of optimal accelerometer wear time to reflect a day's physical activity in a university community and in the NHANES study.

Project One

Evaluation of the MyWellness Key Accelerometer.

The first project described a small study designed to evaluate a commercial accelerometer for the Technogym company. The accelerometer is called the MyWellness Key and was developed for retail sale to track physical activity in the general population.

The validation study included 16 men and women, ages 20 to 60 years, stratified by sex and activity level (2=Low, 4=Middle, 2=High) using the Short Telephone Activity Recall (Matthews, Ainsworth, Hanby, et al., 2005)(Bouchard et al., 1983). Eight men and eight women were selected using randomized sampling with replacement from a cohort of forty-one volunteers (n=27 women; n=14 men). Volunteers were faculty and staff recruited via flyers and through direct email sent to employees of a university campus (n=500).

The MyWellness Key accelerometer was evaluated for consistency of response during a graded treadmill walking test and during one week of free-living observation. Validation instruments included the ActiGraph GT1M accelerometer, Yamax Digiwalker SW-200 pedometer, the 3-day Bouchard Activity Record (BAR) (Bouchard et al., 1983), and Global Physical Activity Questionnaire (GPAQ) (Armstrong & Fiona Bull, 2006).

Descriptive statistics were computed and presented as median (25th, 75th percentile) for time spent in light, moderate, and vigorous intensity physical activity (Ekelund et al., 2006; Freedson, Melanson, & Sirard, 1998). Spearman rank-order correlation coefficients were plotted and used to investigate the validity of the MyWellness Key with the Actigraph and Digiwalker to measure physical activity in controlled laboratory settings, assessed by calibrated treadmill walking performances. Intensity-specific (light, moderate and vigorous) Spearman rank-order correlations between the MyWellness Key, ActiGraph, BAR and GPAQ (moderate and vigorous only) were used to examine associations between physical activity scores in free-living settings. All p values were two-tailed, and values of less than 0.05 were considered to indicate statistical significance. All statistical procedures were performed by using SAS software (version 9.2; SAS Institute, Cary, North Carolina).

Results from the study were presented in abstract form at the 2009 Southwest Chapter of the American College of Sports Medicine 29th Annual Meeting and published in the British Journal of Sports Medicine (Herrmann et al., 2009). Project one is presented in Chapter 4.

Project Two

Validity and Reliability of the Global Physical Activity Questionnaire (GPAQ) in Adults.

The second project was performed as an ancillary study to the ASUKI Step worksite walking intervention conducted at Arizona State University and the Karolinska Institute in Sweden. The study was designed to assess the validity and reliability of the Global Physical Activity Questionnaire (GPAQ). The GPAQ is currently used by the World Health Organization for international surveillance of physical activity and inactivity and has not been evaluated fully against complementary physical activity questionnaires and measures of physical fitness and anthropometry.

This study consisted of two sub-studies to evaluate evidence of validity (Study 1) and test-retest reliability (Study 2) for the GPAQ. In Study 1, during a scheduled laboratory visit, 69 subjects already enrolled in the ASUKI Step physical fitness study, completed tests to obtain their height, weight, percent body fat, waist circumference, resting heart rate, resting blood pressure, and estimated VO_2 max. At the end of the testing session, participants were given an ActiGraph GT1M accelerometer to wear for the next seven days. Approximately 3-5 days after the laboratory visit, participants were called by a study staff member to complete the GPAQ. All study participants were instructed to complete a web-based survey which included the short IPAQ. Pedometer data were obtained from the computerized, password-protected program accessed by participants to record their daily steps.

Correlations were computed between the GPAQ and two objective measures of physical activity (pedometer and accelerometer), a subjective measure of physical activity (IPAQ), and measures of physical fitness (estimated VO_2 max), body fatness (percent fat, waist circumference, and BMI), and cardiovascular health (blood pressure and resting heart rate). A multivariate analysis of variance with the least significant difference (LSD) post hoc method was performed to assess differences in the measures between the three GPAQ categories (low, moderate, and high).

Weighted Cohen’s kappa coefficients and percent agreement were used to compare categorical scores from the GPAQ and IPAQ for low, moderate, and high physical activity levels (Cohen, 1992). Bland-Altman plots were constructed to compare the assessment of MVPA minutes of the GPAQ, IPAQ and ActiGraph accelerometer data.

Study 2 data were used to assess test-retest reliability evidence for the GPAQ physical activity scores. Study 2 was a two-week, cross-sectional study designed to evaluate a new commercial accelerometer described earlier as Project 1 (Herrmann, Hart, Lee, & Ainsworth, 2009). Sixteen participants (n=8 men, n=8 women) were randomly selected from 41 volunteers for an accelerometer validation study. The GPAQ test-retest reliability data were obtained by telephone interview, 10 days apart.

Test-retest reliability of the GPAQ scores was computed by intensity level and domain-specific physical activity using intra-class correlation coefficients (ICC).

Results from the project were presented in abstract form at the 2010 International Congress for Physical Activity and Health and at the 2009 Southwest Chapter of the American College of Sports Medicine 29th Annual Meeting. The combined results of Study 1 and Study 2 have been submitted for publication in a peer reviewed journal. Project two is presented in Chapter 5.

Project Three

How Many Hours are Enough? Optimal Accelerometer Wear Time to Reflect Daily Physical Activity

The third project is a secondary study to the ASUKI Step worksite walking intervention conducted at Arizona State University and the Karolinska Institute in Sweden. The purpose was to understand the influence of daily accelerometer wear time on physical activity data.

Accelerometer data were obtained from 124 study participants randomly selected to participate in the ASUKI Step physical fitness substudy. Collectively, the 124 study participants wearing the accelerometers contributed approximately 1,200 days of accelerometer monitoring data. The accelerometers were scored to assess time spent in each intensity level (inactivity, light, moderate, and vigorous) using a SAS statistical program. Then the data were split into two samples (original and validation sample) to compare the accuracy of the modeling performed with the original accelerometer data sample.

The criterion day data set, used in the semi-simulated approach, was set at 14 hours per day of accelerometer wear time. Additional data sets of 13-, 12-, 11-, and 10 hours per day of accelerometer wear time were then selected to use as reference for missing data. The semi-simulation approach removed data from the known 14 hour criterion day data set by matching the missing data pattern of the 13-, 12-, 11-, and 10 hour data sets. This was done to allow a comparison of the $min \cdot d^{-1}$ spent in activity at varying intensities by different hours of wear time.

To assess differences in daily minutes between the semi-simulation data sets (13-, 12-, 11-, and 10 hours per day) and criterion day (14 hours per day) at each intensity level (inactivity, light, moderate, and vigorous), repeated measures ANOVAs with the Least Significant Difference (LSD) post hoc method and Absolute Percent Error (APE) were used. APE identified the percent difference between two values with a lower APE desired (Kang et al., 2009). Analyses were performed on both the original and validation data sets.

Results from this project were presented in abstract form at the 2010 American College of Sports Medicine's 57th Annual Meeting and have been submitted for publication in a peer reviewed journal. Project three is presented in Chapter 6.

Project Four

Validity of Movement and Activity in Physical Space (MAPS) Scores in Healthy Adults.

The purpose of project four was to examine the characteristics of MAPS scores measured by a combination of data from global positioning systems receivers and accelerometry in adults. This study was a cross-sectional design to evaluate the validity and reliability of MAPS scores in 75 healthy adults.

Participants completed three days of monitoring while wearing an accelerometer (measure of physical activity) and a GPS receiver (measure of environment) on a waist belt during all waking hours and completed a detailed travel diary. Physical activity and environmental data were combined using the Movement and Activity in Physical Space (MAPS) score. The MAPS formula was created to incorporate measures of activity, time, and location data from GPS, geographic information systems (GIS) and accelerometers to produce a single composite score (Herrmann, Ragan, et al., 2008). A higher MAPS score indicates a higher level of free-living function which is characterized by a combination of activity and environmental interaction.

Descriptive statistics of study participants were computed and presented as mean and standard deviation by sex and across tertiles of MAPS scores. The relationship between MAPS and other physical activity related measures (GPAQ, Social Support for exercise, and Self-efficacy for exercise) was evaluated using a Pearson Product Moment Correlation ($\alpha = .05$). The reliability of MAPS scores was evaluated using an Intraclass Correlation Coefficient (ICC). All analyses were performed in SAS 9.2 with alpha level 0.05.

Preliminary results from this project were presented in abstract form at the 2010 Southwest Chapter of the American College of Sports Medicine 30th Annual Meeting. A manuscript will be submitted for peer-reviewed publication. Project four is presented in chapter 7.

Project Five

Impact of Accelerometer Wear Time on Physical Activity Data.

The purpose of project five was to evaluate the effects of varying amounts of daily accelerometer wear time on physical activity in the 2005-2006 National Health and Nutrition Examination Study (NHANES). This project, an extension of project three, helped to further explain the recommendation for 13-hours per day of accelerometer wear time for a valid assessment of daily activities levels at varying intensities using a larger, nationally representative sample.

NHANES is an ongoing study that employs a complex, multistage probability sampling method to obtain a representative sample of the U.S. population. The purpose of NHANES is to assess the health and nutritional status of adults and children in the United States for use in understanding the prevalence and risk factors for diseases. NHANES participants undergo extensive evaluations that include interviews and health examinations. In 2003, as part of the evaluation process, all ambulatory participants >6 years were asked to wear an accelerometer. The raw data file for NHANES 2005-2006 was released in June 2008.

The full 2005-2006 Physical Activity Monitor data set was downloaded from the NHANES website (http://www.cdc.gov/nchs/nhanes/nhanes2005-2006/exam05_06.htm). NHANES data from adults 18 - 65 years will be selected for use in the data analysis. A SAS statistical program was used to identify outlier values due to accelerometer malfunction. This program was used to identify non-wear periods and time spent in each intensity level (inactivity, light, moderate, and vigorous) for use in the semi-simulation approach.

The criterion wear time value (criterion day) was set at 14 hours per day. A criterion day dataset of 200 days were randomly selected from participants that wore the accelerometer for 14 valid hours. Additional 200-day samples of 13-, 12-, 11-, and 10 hours per day of accelerometer wear time were then be selected and used as reference for missing data in the semi-simulation approach.

The semi-simulation approach (as used in project 3) removed data from the known 14 hour criterion day data set by matching (in a one day-to-one day comparison) the missing data pattern of the 13-, 12-, 11-, and 10 hour data sets. This was done to allow a comparison of the $\text{min}\cdot\text{d}^{-1}$ spent in activity at varying intensities by different hours of wear time. This procedure was repeated to create data sets of 12-, 11-, and 10 hours.

Repeated measures ANOVAs with the Least Significant Difference (LSD) post hoc method was performed to assess differences in daily minutes between semi-simulation data sets (13-, 12-, 11-, and 10 hours per day) and criterion day (14 hours per day) at each intensity level: inactivity, light, moderate, and vigorous. APE was computed between the criterion day and each of the semi-simulation data sets for daily minutes of inactivity, light, moderate, and vigorous activity.

The results from this project will be submitted for presentation at the 2012 ACSM annual meeting and submitted for peer-reviewed publication. Project 5 is presented in chapter 8.

Chapter 4

EVALUATION OF THE MYWELLNESS KEY ACCELEROMETER

Herrmann SD, Hart TL, Lee CD, Ainsworth BE. Evaluation of the Mywellness Key Accelerometer. *British Journal of Sports Medicine,* 2011;45:109–113.

Abstract

Purpose. To examine the concurrent validity of the Technogym MyWellness Key accelerometer against objective and subjective physical activity measures.

Design. Randomised, cross-sectional design with two phases. The laboratory phase compared the MyWellness Key with the ActiGraph GT1M and the Yamax SW200 Digiwalker pedometer during graded treadmill walking, increasing speed each minute. The free-living phase compared the MyWellness Key with the ActiGraph, Digiwalker, Bouchard Activity cord (BAR) and Global Physical Activity Questionnaire (GPAQ) for seven continuous days. Data were analysed using Spearman rank-order correlation coefficients for all comparisons.

Setting. Laboratory and free-living phases.

Participants. Sixteen participants randomly stratified from 41 eligible respondents by sex (n=8 men; n=8 women) and physical activity levels (n=4 low, n=8 middle and n=4 high active).

Results. There was a strong association between the MyWellness Key and the ActiGraph accelerometer during controlled graded treadmill walking ($r=0.91,p<0.01$) and in free-living settings ($r=0.73$–0.76 for light to vigorous physical activity, respectively, $p<0.01$). No associations were observed between the MyWellness Key and the BAR and GPAQ ($p>0.05$).

Conclusions. The MyWellness Key has a high concurrent validity with the ActiGraph accelerometer to detect physical activity in both controlled laboratory and free-living settings.

Introduction

Since 2000, there have been nearly 1200 articles cited in the PubMed database that describe the use of accelerometers in physical activity research. Accelerometers are valued as an objective measure in physical activity research as they have the ability to measure and record information about the duration, intensity and frequency of human movement (Welk, Blair, Wood, et al., 2000). Accelerometers also have been used for public health surveillance in the USA (Matthews et al., 2008; Troiano et al., 2007) to provide objective data about the proportion of adults who meet national physical activity recommendations (Haskell et. al., 2007; Pate et al., 1995; USDHHS, 2008). There is a growing recognition that accelerometers are useful to monitor physical activity and to provide feedback to community residents interested in improving their physical activity, physical fitness and health (Keyserling, Hodge, Jilcott, et al., 2008). A wide range of different accelerometers have been developed to monitor physical activity for research (e.g., ActiGraph), estimate energy expenditure (e.g., SenseWear Pro, Kenz Lifecorder EX) and to provide feedback about physical activity goals (e.g., Philips New Lifestyles). However, there are few methods for the general public to objectively assess the amount of time they spend in moderate to vigorous physical activity. The value of such accelerometers is that they allow users to view and monitor graphical depictions of their activity for tracking activity progression and for motivational purposes. Accordingly, there is much interest in determining the accuracy of accelerometers (Matthews, 2005; Hendelmen, Miller, Baggett, Debold, & Freedson, 2000; Swartz et al., 2000; Welk, 2005; Kayes, Shulter, McPhearson, et al., 2009). Studies that have evaluated accelerometers against criterion measures (e.g., direct observation, indirect calorimetry) suggest they have acceptable validity to measure physical activity (Welk, et al., 2000) and estimate energy expenditure (Welk, Blair, Wood, Jones, & Raymond, 2000). Technogym, known widely for their gym and fitness equipment, has recently developed a new accelerometer called the MyWellness Key designed for use by the general public and in fitness settings to detect the intensity and duration of movement and to provide feedback to the user via an interactive, web-based system. In this paper, we investigate the concurrent validity of the Technogym MyWellness Key accelerometer against objective and subjective physical activity measures in 16 healthy men and women with varying physical activity levels.

Methods

Participants

To provide a comprehensive review of the accelerometer, recruitment targeted men and women, ages 20–60, with physical activity levels ranging from inactive to highly active. The goal was to enroll 16 adults, stratified by sex and activity level, using randomized sampling with replacement methods from a cohort of study volunteers. Volunteers were recruited via flyers and through direct email sent to employees of a university campus (n=500). Forty-one volunteers (n=27 women; n=14 men) responded to the advertisements. To establish eligibility, volunteers completed the three-item, Short Telephone Activity Recall (STAR) (Matthews, Ainsworth, Hanby, et al., 2005) and the Physical Activity Readiness Questionnaire (PAR-Q) (Thomas, Reading, & Shephard, 1992), respectively, to identify physical activity levels and contraindications for exercise. Sixteen study participants were selected randomly from volunteers with equal representation by sex (eight men and eight women) and physical activity level (low n=4, middle n=8 and high active n=4). Prior to completing any study activities, participants read and signed an informed consent form approved by the Institutional Review Board at Arizona State University.

Instruments

The data collection instruments are described below.

Short Telephone Activity Recall. The STAR (Matthews et al., 2005) categorises physical activity levels as low, middle and high according to the CDC-ACSM recommendations (Haskell, et al., 2007; Pate, et al., 1995). The STAR has three questions, is self administered and takes about one minute to complete. It has a high validity against repeat 24-h physical activity records (r=0.91) and a moderate 1-week test–retest reliability (r=0.55). The sensitivity and specificity to categorise adults into physical activity categories is 0.50 and 0.65, respectively (Matthews, et al., 2005).

Physical Activity Readiness Questionnaire. The PARQ (Thomas, et al., 1992) is a seven-item, self-administered questionnaire designed to identify participant's with contraindications to exercise. The PAR-Q is self administered and takes less than one minute to complete.

Bouchard Activity Record. The Bouchard Activity Record (Bouchard, Tremblay, Leblanc, et al., 1983) is a self-report, 3-day physical activity log used to determine the type of physical activity performed during two weekdays and one weekend day. Participants identify the type of activity performed every 15 min using a predetermined checklist with nine activity categories (e.g., 1=lying to 9=high-intensity sport activities) ranging from an energy cost from 1.0 to 7.8 METs. The BAR is self administered and has a concurrent validity of 0.31 with cycle ergometer test for physical work capacity with a high test–retest reliability (intraclass correlation coefficient=0.97) (Bouchard et al., 1983).

Global Physical Activity Questionnaire. The Global Physical Activity Questionnaire (GPAQ) is a 1-week, telephone administered, recall questionnaire used to asses physical activity in the World Health

Organisation STEPS global surveillance activities (Armstrong & Bull, 2006). The GPAQ identifies physical activity at work, travel to and from places, recreational activities and time spent in inactivity. Intensity levels are moderate (4 METs), vigorous (8 METs) and inactive (1 MET). The GPAQ takes about five minutes to complete and is scored by multiplying the minutes per week for each activity by their associated metabolic equivalents (METs) to create MET-min scores. Activity specific scores are summed to create total MET-min/week. Test–retest reliability for the GPAQ ranges from (r=0.67–0.81), and validity for total activity versus pedometers is r=0.31 (Armstrong & Bull, 2006).

ActiGraph GT1M. The ActiGraph model GT1M accelerometer (ActiGraph, LLC, Pensacola, Florida, USA) is a uniaxial piezoelectric accelerometer (3.8×3.7×1.8 cm; 27 g) that assesses physical activity intensity, duration, steps and an estimate of physical activity caloric expenditure. The GT1M records vertical accelerations from approximately 0.05*g* to 2.0*g* with a frequency response from 0.25 to 2.50 Hz. Output data are digitised at a rate of 30 times per second with intensity data recorded in epochs (sampling interval). Consistent with current practice, the epoch was set at one minute. Data are downloaded using ActiGraph software, converted to activity counts and stored in a computer database.

Yamax Digiwalker SW-200. The Yamax Digiwalker SW-200 (Yamax Corporation, Tokyo, Japan) is an electronic pedometer that has been widely used in behavioural research and to support behaviour change (Tudor-Locke, McClain, Abraham, et al., 2009). The SW-200 measures vertical accelerations at the hip by a spring-suspended horizontal lever arm. The up and down movement of the horizontal lever arm opens and closes an electrical circuit which causes the device to register each step. The Digiwalker provides a single output of steps on a digital display screen and is currently considered the “criterion pedometer” for measurement of steps taken (Tudor-Locke et al., 2009). The SW-200 is highly accurate and has demonstrated the ability to measure step counts within 1% of actual steps (Schneider et al., 2004).

Technogym MyWellness Key. The MyWellness Key (Technogym, Gambettola, Italy) is a new single axis accelerometer intended for use by the general population for monitoring their physical activity (figure 1). The MyWellness Key is a small, lightweight (8.5×2.0×0.7 cm; 18.7 g) monitor that is worn on the waist band to measure activity with a simple user-friendly unit (MOVE score) and time spent in different intensity levels (light: 1.8–2.9 METs, moderate: 3–5.9 METs and vigorous: 6+ METs). Intensity levels are determined using proprietary web-based software. The MyWellness Key has a sampling frequency of 16 Hz and can detect acceleration ranging in magnitude from 0.06*g* to 12.0*g* with a frequency response from 0.1 to 5 Hz. The MyWellness Key has a USB connector used to recharge the battery and interface with a web-based program to download, score and record a user's physical activity data. The battery has sufficient charge to record data for 49–59 days depending on the type of usage. The MyWellness Key uses an algorithm to automatically see the user's daily goal based on the previous seven days of recorded activity. The user goal is indicated by a white bar that the user has to fulfill during the day until a "+" appears on the display screen. The web-based program also can be used for physical activity goal setting as the website offers a variety of tailored exercise programs, discussion boards, daily encouragement and physical activity recommendations. The MyWellness Key also has an added feature of communicating with specific fitness equipment being used by the participant in order to better assess physical activity. This feature was not tested in this study.

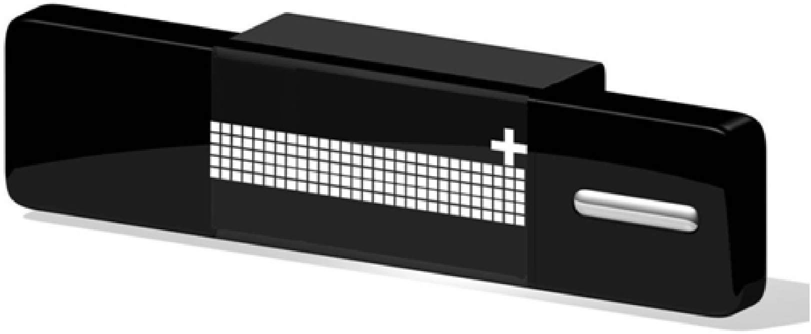


Figure 1. The MyWellness Key accelerometer with the display screen showing the user daily goal has been achieved.

Measurement procedures

The procedures for this study were approved by Arizona State University Institutional Review Board. The study was divided into a controlled laboratory and a free-living phase. Participants were paid $50 for successfully completing all study procedures.

Laboratory phase. The MyWellness Key was compared with the ActiGraph and the Digiwalker during graded treadmill exercise tests in a controlled environment. Weight and height were measured before being fitted with an elastic waist belt with the motion sensors attached to wear during treadmill walking. The ActiGraph was worn on the left mid-axillary line of the hip, and the MyWellness Key was positioned in line with the midline of the right anterior thigh. The Digiwalker was worn over the midline of the left anterior thigh. Data from each instrument were recorded throughout the test. The treadmill protocol included eight 1-min stages to evaluate the ability of the MyWellness Key to detect varying walking speeds ranging from slow to very fast. The treadmill grade remained at 0% throughout the test. The speed started at 53.6 m/min and increased by 8.1 m/min until reaching 110.0 m/min. Each stage was followed by a one minute clearing period to record Digiwalker steps. Data from the ActiGraph and MyWellness Key were downloaded onto a PC for later processing and analysis.

Free-living phase. The MyWellness Key was compared with the ActiGraph, Digiwalker, BAR and the GPAQ during usual daily activities. Participants wore an elastic belt holding the three monitors for seven continuous days. Upon waking, participants attached the belt (or immediately after bathing or showering), wore the monitors all day (except in water) and removed it before going to sleep at night. Because the Digiwalker has no memory function, participants recorded their steps every evening on a steps recording form and reset the Digiwalker to zero each morning before engaging in their daily activities. The BAR was completed for one weekend day and two weekdays to obtain a self-report of activity performed while wearing the monitors. The weekend day and weekdays were not standardised; thus, the days with data recorded varied. Participants were administered the GPAQ on day five to determine agreement between the MyWellness Key's physical activity scores with self-reported physical activity levels. Participants returned all monitors to the study centre following the free-living phase.

Data Treatment and Statistical Analysis

The accelerometers were downloaded per manufacturers' instructions. To determine aggregate daily minutes for the MyWellness Key, data were recorded from the proprietary MyWellness Key web-based software for time in each intensity level. For the ActiGraph, time spent in each intensity level was determined using a SAS Statistical Program.

Matthews' cutpoints were used to identify time spent in light intensity (100–1951 cts/min)(Matthews, 2005) and Freedson's cutpoints were used to determine time spent in moderate (1952–5724 cts/min) and vigorous intensity physical activity (5725+ cts/min) (Freedson, Melanson, & Sirard, 1998). Minutes spent in intensity levels were averaged across seven days. For the ActiGraph, 60 consecutive minutes with no movement data were considered to be non-wear time, and only days with ≥10 hours of wear time were included in the analysis (Troiano, et al., 2008). Daily steps obtained from the Digiwalker were averaged across the seven days. Descriptive statistics of study participants were computed and presented as median (25th, 75th percentile). The Shapiro–Wilk's tests and normal probability plots were used to test normality assumptions. Spearman rank-order correlation coefficients were plotted and used to investigate the concurrent validity of the MyWellness Key with the Actigraph and Digiwalker to measure physical activity in controlled laboratory settings, assessed by calibrated treadmill walking performances. Intensity-specific (light, moderate and vigorous) Spearman rank-order correlations between the MyWellness Key, ActiGraph, BAR and GPAQ (moderate and vigorous only) were used to examine associations between physical activity scores in free-living settings. Partial Spearman correlations were also tested in all analyses after adjustment of age, body mass index (BMI) and sex. All p-values were two-tailed, and values of less than 0.05 were considered to indicate statistical significance. All statistical procedures were performed by using SAS software (version 9.2; SAS Institute, Cary, North Carolina).

Results

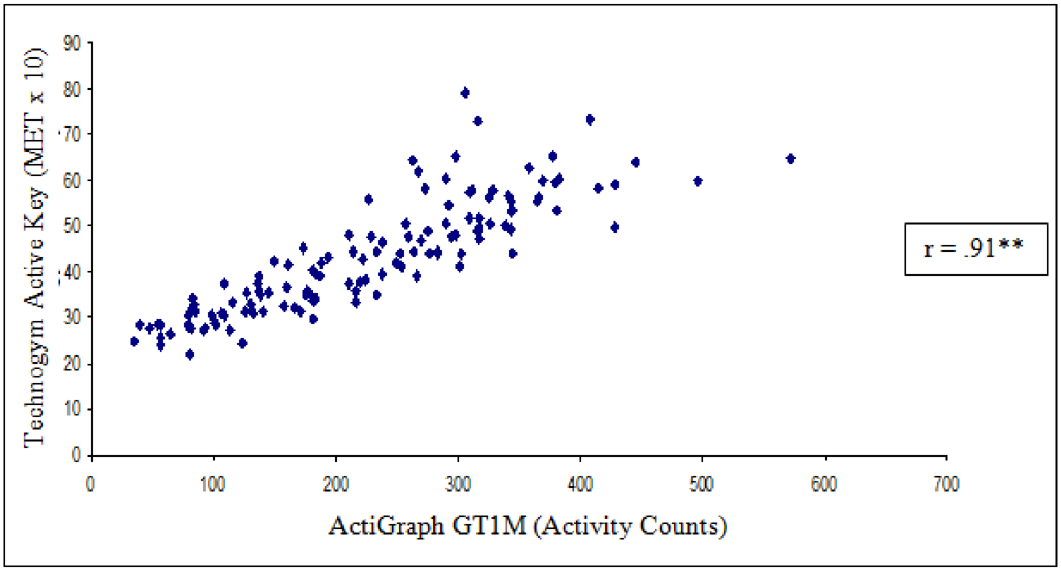
The descriptive characteristics of the study participants are presented in table 1. Participants were approximately 40 years old with height and weight reflecting normal BMI levels. As intended, participants ranged between low to high levels of physical activity.

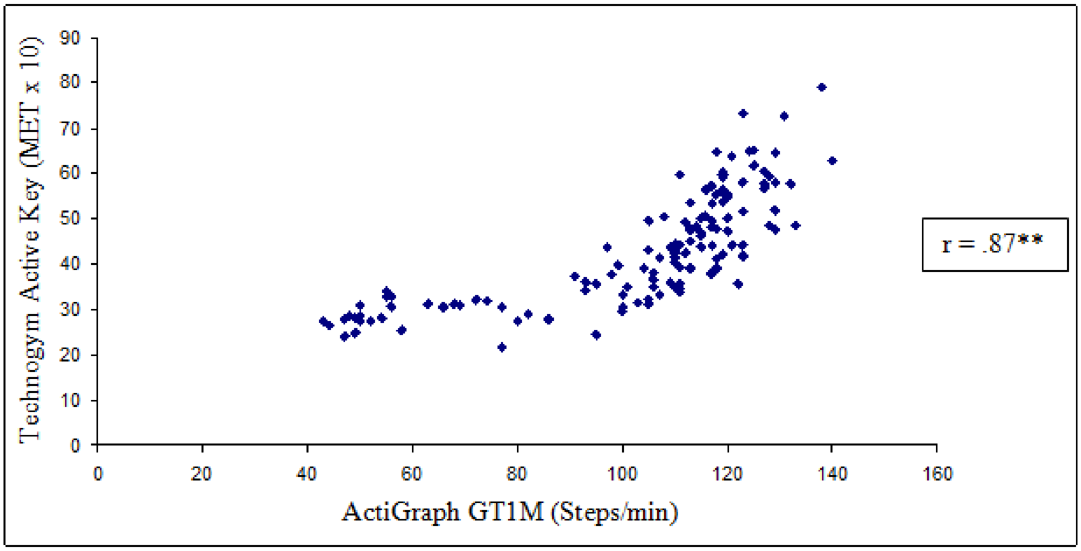
Table 1

Descriptive Characteristics of the Study Sample (N=16)

Variable	Total	Men (n=8)	Women (n=8)
Age (yr)	40.2 ± 12.6	38.9 ± 12.4	41.5 ± 13.4
Height (cm)	175.8 ± 7.0	190.6 ± 5.6	170.9 ± 4.4
Weight (kg)	77.2 ± 12.3	85.0 ± 10.4	69.4 ± 9.0
BMI (kg/m^2)	25.1 ± 3.4	26.0 ± 3.1	24.2 ± 3.6

Figure 2 shows a plot of the results from the laboratory phase for the comparison of the MyWellness Key accelerometry output with the (a) ActiGraph counts, (b) ActiGraph steps and (c) Digiwalker steps during the laboratory phase treadmill test. Strong associations were observed between the MyWellness Key and the ActiGraph activity counts (r=0.91, p<0.01), ActiGraph steps (r=0.87, p<0.01) and Digiwalker steps (r=0.81, p<0.01).





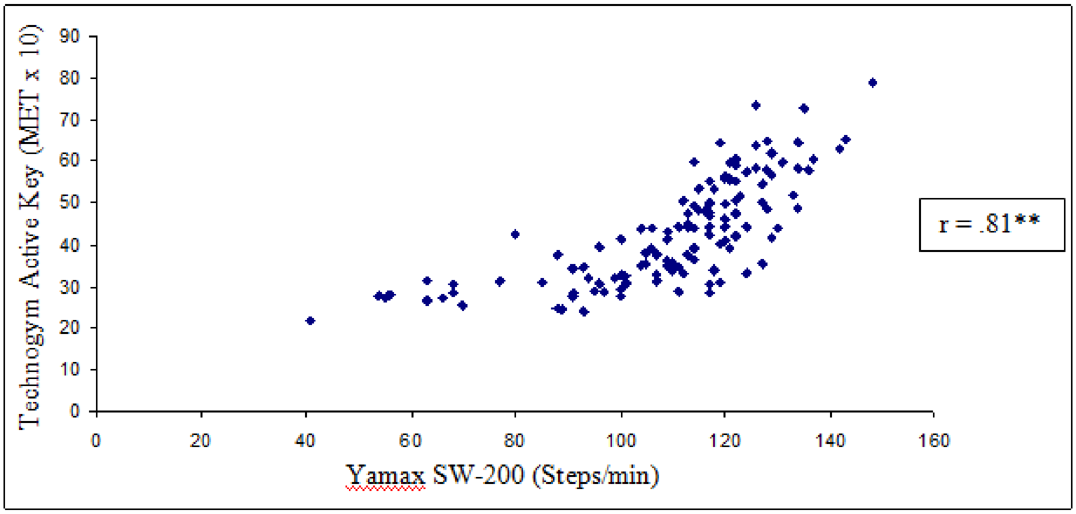


Figure 2. Spearman rank-order correlation coefficient plots between the (A) MyWellness Key and Actigraph activity counts, (B) Actigraph steps/min, and (C) Yamax steps/min during calibrated treadmill walking. **p<0.01.

Table 2 shows the results from the free-living phase of the study. Participants took an average of 9,452 (6,271; range=1,499–21,671) steps per day as detected by the Digiwalker. Average wear time (based on ActiGraph data) was 11.3±1.5 h/day. Correlations were high between the MyWellness Key and ActiGraph (r=0.73–0.76, p<0.05) with median minutes spent in moderate and vigorous physical activity higher for the MyWellness Key than for the ActiGraph moderate (Δ=9 min; p<0.01) and vigorous (Δ=2 min; p=0.68) physical activity. Minutes recorded in light physical activity were higher on the ActiGraph than the MyWellness Key (Δ=149 min; p<0.01). No associations were observed between the MyWellness Key and the BAR or GPAQ (p>0.05). These associations remain unchanged after further adjustment for covariates (age, BMI and sex).

Table 2

Median (25th, 75th Percentile) and Range for Minutes per Day During 7 Days of Free-living Activity and Spearman Rank-order Correlations Between the MyWellness Key with Other Physical Activity Measures

Instrument	Intensity	Median (25th,75th %) (min)	Range (min)	Active Key (r)
Technogym MyWellness Key[a]	Light	97 (-75,121)	59 – 172	-
	Moderate	33 (26, 51)	12 – 100	-
	Vigorous	2 (0, 5)	0 – 16	-
ActiGraph GT1M[a]	Light *(100-1951 $cts \cdot min^{-1}$)*	237 (188, 315)	111 – 532	0.76**
	Moderate *(1952-5724 $cts \cdot min^{-1}$)*	24 (10, 50)	2 – 132	0.76**
	Vigorous *(5725+ $cts \cdot min^{-1}$)*	0 (0, 4)	0 – 50	0.73**
Bouchard Activity Record[b]	Light	535 (-331,766)	111 – 1791	0.09
	Moderate	65 (25, 143)	0 – 360	0.34
	Vigorous	0 (0, 15)	0 – 70	0.15
GPAQ[c]	Moderate	60 (30, 135)	10 – 280	-0.03
	Vigorous	28 (0, 60)	0 – 90	0.40

[a] Light = 1.8-2.9 METs; Moderate = 3-5.9 METs; Vigorous = 6+ METs
[b] Light = 2.3-2.9 METs; Moderate = 3-5.9 METs; Vigorous = 6+ METs
[c] Global Physical Activity Questionnaire; Moderate = 4.0-7.9 METs; Vigorous = 8+ METs
* p < .05; ** p < .01

Discussion

This study was performed to determine the concurrent validity of the MyWellness Key accelerometer developed for use by exercise practitioners and the general population. Comparison of physical activity measured during treadmill walking showed a high association between the MyWellness Key and the ActiGraph during slow to fast walking ($r=0.91$). Correlations also were moderate-to-high during a 7-day free-living period ($r=0.73–0.76$, $p<0.05$). This suggests high concurrent validity between the MyWellness Key and the ActiGraph. Correlations were lower and not statistically significant between the MyWellness Key and the GPAQ and BAR. Differences in the median minutes of light and moderate intensity physical activity for the MyWellness Key and the ActiGraph may be attributed to cutpoints used to assign minutes at each intensity level, the sensitivity of the instruments to detect motion at varying movement speeds and the variability in the self-report versus objective physical activity. In 2007, Ham et al. showed that large differences in the minutes reported on accelerometers are largely due to cutpoints used to denote intensity levels. Using data from heart rate monitors to determine physiological intensity and the ActiGraph to assess movement, when comparing time in moderate physical activity, values ranged from 17.9 to 139 min/day (Ham, Reis, Strath, et al., 2007). The Freedson cutpoint (1951 cts/min) resulted in the least time of moderate physical activity, and the Hendelman cutpoint (191 cts/min) resulted in the most time in moderate physical activity. Despite these differences in the absolute minutes, high correlations between the MyWellness Key and the ActiGraph during graded treadmill walking and in free-living activities suggest that the instruments are measuring similar increases in the energy cost of physical activity.

The strong association between the MyWellness Key and ActiGraph during increased treadmill walking speeds suggests that the MyWellness Key is sensitive in detecting increasing walking intensity. However, when the MyWellness Key was compared with step counts from the ActiGraph and Digiwalker, there was a flat response by the MyWellness Key at slower walking speeds (56–78 m/min). As depicted in figure 2B, ActiGraph steps increased with increasing speeds, but the MyWellness Key output changed only at speeds >90 steps/min. This is likely related to the MyWellness Key reflecting only the MET intensity of physical activity at slower speeds and not steps taken as with the ActiGraph and the Digiwalker. Hence, the MyWellness Key output does not change until the speed reached a moderate energy cost. Another possible reason for discrepancy in output detected during the treadmill test may be due to the MyWellness Key’s lower acceleration detection capability (0.06*g*) than the Digiwalker (0.35*g*). Consistent with the ActiGraph, during treadmill speeds requiring >90 steps/min, association between the MyWellness Key and Digiwalker increased linearly.

Interest in the ActiGraph to assess daily physical activity has led to the development of various cutpoints to estimate time in intensity levels. We used the Matthews et al (2005) and Freedson et al (1998) ActiGraph cutpoints and found high correlations with the Freedson cutpoints for moderate and vigorous physical activity and with Matthews' cutpoint for light activity with the MyWellness Key (r=0.73–0.76). Comparisons between the MyWellness Key and BAR were not statistically significant. However, the correlation between the moderate physical activity scores was r=0.34, which is similar to findings reported between the BAR and accelerometers (Schmidt, Freedson, & Chasan-Taber, 2003). The BAR collects data in 15-min increments, whereas the MyWellness Key tracks activity each minute. This difference may account for the lower correlations. The BAR also detects different types of physical activity that the MyWellness Key cannot detect (eg, sitting and standing), leading to more minutes of light physical activity. Similar to the BAR, comparison of the MyWellness Key and the GPAQ yielded a correlation of r=0.40 for vigorous activity ($p>0.05$). This correlation is similar to those reported for validation of physical activity questionnaires using accelerometers (Boon, Hamlin, Steel, et al., 2010).

The value of this validation study is that it demonstrates the accuracy of a new commercial accelerometer developed for use in community and fitness club settings to track daily physical activity. While the sample size was small for a validation study, the participants were diverse in their physical activity levels from inactive to very active and equal in number by sex. Further, compliance was high with no drop-outs and all participants completing each study task.

In summary, the MyWellness Key accelerometer is designed for use in the general public to track physical activity using a web-based interface demonstrated acceptable concurrent validity with objective measures of physical activity. There was a strong association between the MyWellness Key accelerometer and the ActiGraph GT1M accelerometer outputs during controlled laboratory and in free- living settings.

References

Armstrong, T. & Bull F. (2006) Development of the World Health Organization Global Physical Activity Questionnaire (GPAQ). *Journal of Public Health, 14,* 66–70.

Boon, R. M., Hamlin, M. J., Steel, G. D., & Ross, J. J. (2010). Validation of the New Zealand Physical Activity Questionnaire (NZPAQ-LF) and the International Physical Activity Questionnaire (IPAQ-LF) with Accelerometry. *British Journal of Sports Medicine, 44, 741-746*.

Bouchard, C., Tremblay, A., Leblanc, C., Lortie, G., Savard, R., & Thériault, G. (1983). A method to assess energy expenditure in children and adults. *American Journal of Clinical Nutrition, 37*, 461-467.

Crouter, S. E., Schneider, P., Karabulut, M., & Bassett, D. (2003). Validity of 10 Electronic Pedometers for Measuring Steps, Distance, and Energy Cost. *Medicine & Science in Sports & Exercise, 35*(8):1455-1460.

Freedson, P. S., Melanson, E., & Sirard, J. (1998). Calibration of the Computer Science and Applications, Inc. accelerometer. *Medicine & Science in Sports & Exercise, 30,* 777-781.

Ham, S. A., Reis, J. P., Strath, S. J., Dubose, K. D., & Ainsworth, B. E. (2007). Discrepancies between methods of identifying objectively determined physical activity. *Medicine & Science in Sports & Exercise, 39*, 52-58.

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. *Medicine and Science in Sports and Exercise*, *39*,1423-34.

Hendelman, D., Miller, K., Baggett, C., Debold, E., & Freedson, P. (2000) Validity of accelerometry for the assessment of moderate intensity physical activity in the field. *Medicine and Science in Sports and Exercise,32*, 442–449.

Kayes, N. M., Schulter, P. J., McPherson, K. M., Leete, M., Mawston, G., & Taylor, D. (2009). Exploring Actical accelerometers as an objective measure of physical activity in people with multiple sclerosis. *Archives of Physical Medicine & Rehabilitation, 90*, 594-601.

Keyserling T. C., Hodge C. D., Jilcott S. B., Johnston, L. F., Garcia, B. A., Gizlice, Z., et al. (2008). Randomized trial of a clinic-based, community-supported, lifestyle intervention to improve physical activity and diet: The North Carolina enhanced WISEWOMAN project. *Preventive Medicine, 46,* 499-510.

Matthews, C. E. (2005). Calibration of accelerometer output for adults. *Medicine and Science in Sports and Exercise, 37*, S512-522.

Matthews, C. E., Ainsworth, B. E., Hanby, C., Pate, R. R., Addy C., Freedson P. S., *et al.* (2005). Development and testing of a short physical activity recall questionnaire. *Medicine & Science in Sports & Exercise, 37*, 986-994.

Matthews, C. E., Chen, K. Y., Freedson, P. S., Buchowski, M. S., Beech, B. M., Pate, R. R., Troiano R. P. (2008). Amount of time spent in sedentary behaviors in the United States, 2003-2004. *American Journal of Epidemiology*, *167*, 875-881.

Pate, R. R., Pratt, M., Blair, S. N., Haskell, W. L., Macera, C. A., Bouchard, C., et al. (1999). Physical activity and public health. A recommendation from the Centers for Disease Control and Prevention and the American College of Sports Medicine. *JAMA, 273*, 402-407.

Schmidt, M. D., Freedson, P. S., & Chasan-Taber L. (2003). Estimating physical activity using the CSA accelerometer and a physical activity log. *Medicine & Science in Sports & Exercise, 35*, 1605-1611.

Schneider, P. L., Crouter, S. E., & Bassett, D. R. (2004). Pedometer measures of free-living physical activity: comparison of 13 models. *Medicine & Science in Sports & Exercise, 36*, 331-335.

Swartz, A. M., Strath, S. J., Bassett, D. R., O'Brien, W. L., King, G. A., & Ainsworth, B. E. (2000). Estimation of energy expenditure using CSA accelerometers at hip and wrist sites. *Medicine & Science in Sports & Exercise, 32*, S450-456.

Thomas, S., Reading, J., & Shephard, R. J. (1992). Revision of the physical activity readiness questionnaire (PAR-Q). *Canadian Journal of Sports Science, 17,* 338-345.

Troiano, R. P., Berrigan, D., Dodd, K. W., Mâsse, L. C., Tilert, T., &McDowell, M. (2008). Physical activity in the United States measured by accelerometer. *Medicine and Science in Sports and Exercise*, *40*, 181-188.

Tudor-Locke, C., McClain, J. J., Hart, T. L., Sisson, S. B., & Washington, T. L. (2009). Pedometry Methods for Assessing Free-living Youth. *Research Quarterly for Exercise & Sport, 80*(2), 175-184.

U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease

Prevention and Health Promotion. *Physical Activity Guidelines for Americans*. Atlanta, GA: 2008.

Welk, G. J. (2005) Principles of design and analyses for the calibration of accelerometry-based activity monitors. *Medicine and Science in Sports and Exercise,37*, S501-11.

Welk, G. J., Blair, S. N., Wood, K., Jones, S., & Thompson, R. W. (2000). A comparative evaluation of three accelerometry-based physical activity monitors. *Medicine and Science in Sports and Exercise*, *32*, S489-97.

Chapter 5

GLOBAL PHYSICAL ACTIVITY QUESTIONNAIRE (GPAQ): VALIDITY AND RELIABILITY IN ADULTS

Herrmann, S.D., Heumann, K. J., Ainsworth, B. E. Global Physical Activity Questionnaire (GPAQ): Validity and Reliability. (In review).

Abstract

The World Health Organization uses the Global Physical Activity Questionnaire (GPAQ) and International Physical Activity Questionnaire (IPAQ) to determine the prevalence of worldwide physical activity and determine risk for physical inactivity related chronic diseases. Psychometric properties of the IPAQ (validity and reliability) have been studied in greater detail, while little is known about the validity and reliability of the GPAQ in U.S. adults. Two studies were conducted to evaluate the GPAQ for convergent and concurrent validity (Study 1) and for test-retest reliability (Study 2) evidence. In Study 1, 69 participants (n = 12 men, n = 57 women; M age = 43 years, SD = 12) completed the GPAQ, IPAQ and wore a Yamax SW-200 pedometer and ActiGraph GT1M accelerometer for seven days. In Study 2, 16 participants (n = 8 men, n = 8 women; M age = 40 years, SD = 13) completed the GPAQ 10 days apart. The GPAQ minutes of moderate and vigorous activity were correlated with the ActiGraph moderate ($r = 0.28$, $p = 0.04$) and vigorous (r

$= 0.48$, $p < 0.01$) physical activity. Sedentary ($r = 0.51$, $p < 0.01$), moderate-to-vigorous ($r = 0.48$, $p < 0.01$) and vigorous ($r = 0.63$, $p < 0.01$) intensity physical activity assessed by the GPAQ and IPAQ were related. GPAQ moderate-to-vigorous physical activity was related to percent body fat ($r = -0.32$, $p < 0.01$) and estimated VO2 max ($r = 0.26$, $p = 0.04$). Test-retest reliability was high ($r = 0.83$ to 0.96). These data provide preliminary validity and reliability evidence for the GPAQ that support its use to assess physical activity.

Introduction

Developing standard measures to assess physical activity is of great importance to determine benchmarks of and observe fluctuations in the prevalence of physical activity globally. Questionnaires are commonly used to assess physical activity in intervention and surveillance studies because they are relatively inexpensive to construct, administer, and evaluate as compared with objective measures, such as accelerometers. Within the past 10 years, the World Health Organization (WHO) has recognized the importance of physical activity as a health enhancing behavior and has promoted the assessment of physical activity in its global surveillance system (WHO, 2002; Bauman & Craig, 2005). To this end, the WHO has conducted physical activity surveillance using two surveys, the International Physical Activity Questionnaire (IPAQ) (Bauman & Sallis, 2008; Guthold et al., 2008) and the Global Physical Activity Questionnaire version 2 (GPAQ) (Armstrong & Bull, 2006). Developed for use in the WHO Stepwise [global] surveillance system (STEPS), the GPAQ is a recall of specific domains (work, travel, recreational activity, and inactivity) that reflect physical activity behaviors performed in most countries (Armstrong & Bull, 2006).

Questionnaires are used to assess the prevalence of physical activity in population groups and have the potential to influence policy, therefore, they should be as accurate as possible and evaluated for evidence of validity and reliability (Jacobs, et al., 1993). In examining a questionnaire for validity, an overarching concern is whether the items in a questionnaire reflect the construct of physical activity being assessed. The GPAQ and IPAQ measure moderate- and vigorous-intensity physical activity (MVPA), walking, and inactivity as the physical activity constructs of interest. Unfortunately, there is no “gold standard” with which to compare these constructs, and several forms of validity evidence should be established to provide a comprehensive view of the questionnaire’s ability to assess physical activity levels. Evidence for criterion validity often is established using accelerometers to objectively monitor the frequency and duration of movement by intensity levels. Convergent validity evidence may be established by comparing a new questionnaire with other questionnaires or another physical activity assessment method deemed valid for use in desired settings. Face validity evidence is established by viewing the domains or types of physical activity assessed, such as transportation, household, and leisure-time physical activity, and deciding if these meet the types of physical activity one wishes to assess. Evidence for content validity provides a more detailed evaluation of the types of physical activity assessed than merely looking at the instrument to determine if it contains the types of physical activity desired. Validation studies of physical activity questionnaires commonly compare the questionnaire to direct (e.g., accelerometer, pedometer) and indirect (e.g., waist circumference, BMI, blood pressure, etc.) criterion measures (Craig, et al., 2003; Jacobs, et al., 1993; Rutten, Ziemainz, Schena, Stahl, Stiggelbout, et al., 2003).

As well as containing validity evidence, physical activity questionnaires should present consistent results when administered across repeat administrations from several days to a week apart in the absence of a physical activity behavior changes (i.e., test-retest reliability). A lack of reliability in a questionnaire's score or physical activity classification limits the validity evidence. Since there is limited description of the measurement properties of the GPAQ in the published literature (Armstrong & Bull, 2006; Bull, Maslin, & Armstrong, 2009; Trinh, Nguyen, van der Ploeg, Dibley, & Bauman, 2009), the purpose of this investigation was to evaluate evidence for the validity and reliability of the GPAQ in U.S. adults with data from two studies. The first study examined the validity of the GPAQ as a measure of sedentary, moderate, and vigorous physical activity. The second study examined the test-retest reliability of the GPAQ when administered twice, 10 days apart.

Methods

Study Design and Participant Selection

Two separate samples of adults that reflect a broad range of activity levels were used to evaluate the validity (Study 1) and test-retest reliability (Study 2) of the GPAQ. Participants for both studies were recruited from faculty and staff of a Southwest U.S. collegiate campus who read and signed an informed consent approved by the University's Office for Research Integrity and Assurance (IRB) prior to study involvement. Because the study assessments required participants to walk on a treadmill, exclusion criteria for both samples included the inability to walk, current diagnosis of cardiovascular disease, uncontrolled risk factors for cardiovascular disease (e.g., uncontrolled hypertension), signs and symptoms of cardiovascular intolerance to exercise (e.g., dizziness with exertion), and current pregnancy. Exclusion criteria were determined using the PAR-Q (Thomas, et al., 1992). Table 1 shows the characteristics of each study sample.

Table 1

Descriptive Characteristics of the Study Samples

	Study 1[a]			Study 2[b]		
	Total (*n* = 69)	Women (*n* = 57)	Men (*n* = 12)	Total (*n* = 16)	Women (*n* = 8)	Men (*n* = 8)
Age (yr)	43.1 ± 11.4	44.0 ± 11.0	38.4 ± 13.8	40.2 ± 12.6	41.5 ± 13.4	38.9 ± 12.4
Body Fat (%)	32.9 ± 10.1	35.1 ± 8.7	21.8 ± 10.0	n/a	n/a	n/a
Waist Circumference (cm)	89.1 ± 14.4	89.1 ± 14.4	92.9 ± 16.6	n/a	n/a	n/a
BMI (wt kg/ht m^2)	27.2 ± 6.2	27.1 ± 6.2	27.3 ± 6.7	25.1 ± 3.4	24.2 ± 3.6	26.0 ± 3.1

[a] Study 1 examined validity evidence. [b] Study 2 examined test-retest reliability.

Instruments

The Global Physical Activity Questionnaire (GPAQ). The GPAQ is a telephone administered recall questionnaire used to assess physical activity by the World Health Organization in approximately 50 countries (Armstrong & Bull, 2006). The GPAQ is comprised of 16 items that measure physical activity engaged in for work, travel to and from places, recreational activities, and time spent in inactivity. These activities are collectively referred to as domain-specific activities. Activities are classified into three intensity levels; moderate (4 METs), vigorous (8 METs) and inactivity (1 MET). The GPAQ takes about five minutes to administer and can be scored as a continuous or a categorical score (www.who.int/chp/steps/resources/GPAQ_Analysis_Guide.pdf). The continuous score sums the duration of recalled activity and presents the data as $\text{min}\cdot\text{d}^{-1}$ or $\text{min}\cdot\text{wk}^{-1}$ for each physical activity domain or by intensity levels of light, moderate, or vigorous physical activity. An estimated energy expenditure score can be computed by multiplying minutes for each questionnaire item by their respective MET level to create $\text{MET-min}\cdot\text{d}^{-1}$ or $\text{MET-min}\cdot\text{wk}^{-1}$ scores. The categorical score identifies categories of physical activity as Low, Moderate, or High computed using the following criteria. The high category is awarded when an individual does vigorous-intensity activity on > three d/wk and accumulating at least 1,500 MET-min/wk or > seven days of any combination of walking, moderate-intensity, or vigorous-intensity activities achieving at least 3,000 MET-min/wk. The moderate category is awarded when an individual completes three days of vigorous-intensity activity of at least 20 min/d, five days of moderate-intensity activity or walking of > 30 min/d, or five days of any combination of walking, moderate-intensity or vigorous-intensity activities achieving at least 600 MET-min/wk. The low category is awarded if the individual does not meet the criteria for the moderate or high category. The focus of this paper is to examine the reliability and validity of the GPAQ min$\cdot$d^{-1} scores by intensity levels and to assess the ability of the GPAQ to categorize physical activity levels.

Direct Criterion Measures

The short, self-administered IPAQ is a recall questionnaire developed for use in surveillance settings and in research studies (Craig, et al., 2003). The IPAQ consists of seven items regarding the frequency (in days per week) and duration (in minutes per day) of moderate- (4 METs) and vigorous-intensity (8 METs) physical activity, walking (3.3 METs) and time spent sitting (1 MET). The instructions for the moderate- and vigorous-intensity questions ask respondents to consider leisure, transportation, and occupational physical activities in their answer. The short IPAQ can be obtained from the website, http://www.ipaq.ki.se. The short IPAQ may be scored by duration in min$\cdot$d^{-1}, by frequency in days$\cdot$wk^{-1}, or in MET-min/wk by multiplying the MET intensity for the moderate, vigorous, and walking activities by their reported frequency and duration.

A total physical activity score is expressed as MET-min$^{.}$wk^{-1} and computed by summing the MET-min$^{.}$wk^{-1} for each activity. The IPAQ categorical scores are used to express physical activity levels of low, moderate, and high using the same criteria as the GPAQ (provided above).

The ActiGraph model GT1M accelerometer (ActiGraph, LLC, Pensacola, Florida, USA) was used to record physical activity intensity, frequency, and duration. The GT1M is a solid state sensor (micro-electro-mechanical systems) accelerometer (3.8cm x 3.7cm x 1.8cm; 27 grams) that assesses physical activity intensity, duration, and steps. Worn at the waist level, the GT1M records vertical accelerations ranging in magnitude from approximately 0.05 to 2.0G with a frequency response from 0.25 to 2.50 Hz. Output data are digitized at a rate of thirty times per second with intensity data recorded in one minute epochs (sampling interval). Data were downloaded using ActiGraph software and stored in a computer database.

Accelerometers were scored to assess time spent in various physical activity intensity levels (min$^{.}$d^{-1}) using a SAS statistical program. Matthew's cut-points were used to identify time spent in inactivity (0 – 99 cts$^{.}$min^{-1}), light intensity (100-1951 cts$^{.}$min^{-1}) (18) and Freedson's cut points were used to determine time spent in moderate intensity (1952-5724 cts$^{.}$min^{-1}) and vigorous intensity (5725+ cts$^{.}$min^{-1}), and moderate-vigorous intensity (1952+ cts$^{.}$min^{-1}) (Freedson et al., 1998). Every minute spent in each intensity level was averaged across seven days. For the

ActiGraph, 60 consecutive minutes with no movement data were considered to be non-wear time and only days with ≥ 10 hours of wear time for ≥ four days were included in the analysis (Troian et al., 2008).

The Yamax Digiwalker SW-200 (Yamax Corporation, Tokyo, Japan) was used to record steps taken. The SW-200 is an electronic pedometer that has been widely used in behavioral research to assess physical activity levels and to support behavioral change (Tudor-Locke, et al., 2009). The SW-200 measures vertical accelerations at the hip by a spring suspended horizontal lever arm. The up and down movement of the horizontal lever arm opens and closes an electrical circuit which causes the device to register each step. The SW-200 provides a single output of steps on a digital display screen and is considered valid and reliable for measurement of steps taken, demonstrating the ability to measure step counts within 1% of actual steps (Crouter, et al., 2003; Schneider, et al., 2004). Daily steps obtained from the SW-200 were self-reported and recorded online by the participants.

Indirect Criterion Measures

Physical fitness measures included estimated maximal oxygen uptake (estimated VO_2 max), percent body fat, waist circumference, height and weight, and resting blood pressure.

Cardiorespiratory fitness was assessed by the Åstrand submaximal cycle ergometer test. The test determines heart rate (HR) response, and corresponding RPE, to one or more submaximal work rates and the results are used to calculate the predicted VO_2 max. A work rate was chosen based upon the subject's verbal indication of their activity levels: trained or untrained. Trained participants were defined as participating in 30 minutes or more of MVPA 3-7 days per week. Untrained participants were defined as participating in less than 30 minutes of MVPA three days per week. Participant's work rates were 50 watts for the sedentary to 150 watts for the highly trained. During the cycle ergometer test, participants were required to reach a heart rate of $\geq$125 beats$\cdot$min^{-1} but less than 85% of their age predicted max HR which was used with the work rate in watts to estimate VO_2 max.

Percent body fat was measures using a Tanita Scale (Model TBF-300A; Tanita Corporation: Arlington Heights, IL). The Tanita Scale measures total body water, lean body mass and fat mass using bioelectrical impedance analysis technique. The formula used to calculate percent body fat combines the impedance analysis with height, gender, and age information.

Waist circumference was measured in cm three times using a Gulick II 150 cm tape measure at the location of the umbilicus. If the subject was larger than 150 cm an extension tape was used. The average of the two closest measures was used for data analyses.

Resting heart rate and blood pressure were measured in a seated position after 5 minutes of rest. The Omron automated device (Model HEM-711 DLX) was used to record systolic and diastolic blood pressure in mmHg and resting heart rate in beats$\cdot$min^{-1}. This measurement was taken three times with one minute rest between measures and the average of the closest two readings was used.

Height was measured in centimeters three times with participants standing in their bare feet using a Seca portable stadiometer (RoadRod, Hamburg Germany) with the average of the two closest readings used. Weight was measured in kilograms once using a Tanita scale during the same time that percent body fat was measured. Body mass index (BMI) was computed as weight in kg divided by height in m^2.

Setting and Approach

Study 1: Validity. Study 1 data were used to assess the validity of the GPAQ. Study 1 was a six-month, quasi-experimental (no control group or pre-study assessment) health promotion study designed to increase daily walking. The goal of the walking study was to complete 10,000 steps per day as recorded by a pedometer. This study included sixty-nine participants (*n*=12 men, *n*=57 women; age = 43 ± 12) recruited from 152 participants randomly selected to complete physical fitness tests from the overall study sample of 714 employees who volunteered to participate in a worksite health promotion study. All participants completed a PAR-Q prior to enrollment to rule out the presence of conditions that could be exacerbated by regular physical activity. If eligible, participants read and signed a consent form, received a pedometer, and were given access to a password protected website to record their daily steps walked.

Physical fitness was measured during a scheduled laboratory visit where participants completed tests to obtain their height, weight, percent body fat, waist circumference, resting heart rate, resting blood pressure, and estimated VO_2 max. At the end of the testing session, participants were given an ActiGraph GT1M accelerometer to wear for the next seven days. Approximately 3-5 days after the laboratory visit, participants were called by a study staff member to complete the GPAQ. All study participants were instructed to complete a web-based survey which included the short IPAQ. Pedometer data were obtained from the computerized, password-protected program accessed by participants to record their daily steps. Accelerometers were collected following the seven day wear period, downloaded, and scored for data analysis.

Data collection staff were trained by a study coordinator to collect the laboratory measures and telephone questionnaire. Quality assurance for the accuracy and reliability of measures was determined using standardized laboratory and telephone interview training protocols. Data for study 1 were collected during March 2009.

Study 2: Reliability. Study 2 data were used to assess the test-retest reliability of the GPAQ physical activity scores. Study 2 was a two-week, cross-sectional study designed to evaluate a new commercial accelerometer (Herrmann, Hart, Lee, & Ainsworth, 2011). Sixteen participants (*n*=8 men, *n*=8 women; age = 40 ± 13 years) were randomly selected from 41 volunteers for an accelerometer validation study. Participants were selected using a stratified randomization method designed to enroll participants by sex (8 males and 8 females) and by physical activity levels (4 low, 8 middle, and 4 high). Physical activity levels were determined using the Short Telephone Activity Recall (Matthews, et al., 2005). Selected participants read and signed a consent form. The GPAQ test-retest reliability data were obtained by calling each participant twice, 10 days apart. Data for study 2 were collected between January and March 2009.

Statistical analysis

Descriptive statistics of study participants were computed and presented as mean ± standard deviation. The Kolmogorov-Smirnov test and normal-probability plots were used to test normality assumptions. All p-values were two-tailed, and values of less than 0.05 were considered to indicate statistical significance. All statistical procedures were performed by using SAS statistical software (version 9.2; SAS Institute, Cary, NC).

Study 1: Validity Evidence. Correlations were computed between the GPAQ ($min \cdot d^{-1}$), two objective measures of physical activity (ActiGraph accelerometer [$min \cdot d^{-1}$] and Yamax pedometer [$steps \cdot d^{-1}$]), a subjective measure of physical activity (IPAQ), and measures of physical fitness (estimated VO_2 max), body fatness (percent fat, waist circumference, and BMI), and cardiovascular health (blood pressure and resting heart rate). A multivariate analysis of variance with the least significant difference (LSD) post hoc method was performed to assess differences in the direct and indirect criterion measures between the three GPAQ categories (Low, Moderate, High) (see Table 2).

Weighted Cohen's kappa coefficients and percent agreement were used to compare categorical scores from the GPAQ and IPAQ for low, moderate, and high physical activity levels (Cohen 1992). Bland-Altman plots were constructed to compare the assessment of MVPA minutes of the GPAQ, IPAQ and ActiGraph accelerometer data.

Study 2: Reliability. Test-retest reliability of the GPAQ $min \cdot d^{-1}$ scores were computed for each physical activity intensity level (light, moderate, and vigorous intensity) and domain-specific physical activity (inactivity, travel, moderate- and vigorous work, and moderate- and vigorous recreation) using intra-class correlation coefficients (ICC).

Results

Study 1: Validity

Table 2 presents the means and standard deviations for variables used to evaluate the validity of the GPAQ. Although not statistically significant ($p > 0.05$), there were graded increases in physical activity scores across the GPAQ low, moderate and high categories for physical activity assessed by the GPAQ, IPAQ, pedometers, and the accelerometer. With the exception of the IPAQ, inactivity time decreased with increasing GPAQ categories. Indirect criterion measures also provided some graded responses without reaching statistical significance ($p > 0.05$). Anthropometric measures of waist circumference and percent body fat decreased across GPAQ categories and estimated VO_2 max increased across the categories.

Table 2

Mean and Standard Deviation Values for Indicators of Direct and Indirect of Physical Activity Stratified by GPAQ Activity Levels: Study 1 (n=69)

	GPAQ Category: Low (*n* = 14)	GPAQ Category: Moderate (*n* = 39)	GPAQ Category: High (*n* = 16)
Direct Criterion Measures			
Sedentary/Inactive ($min \cdot d^{-1}$)			
GPAQ	547.5 ± 179.2	529.2 ± 173.4	550.7 ± 166.3
IPAQ	512.8 ± 184.4	413.2 ± 167.5	477.7 ± 169.9
ActiGraph	716.1 ± 128.5	658.5 ± 106.2	600.8 ± 88.1

Moderate (min·d^{-1})			
GPAQ	39.4 ± 24.9	61.7 ± 32.0	122.9 ± 77.8
IPAQ	42.8 ± 45.7	57.3 ± 53.7	56.8 ± 51.7
ActiGraph	26.9 ± 17.4	34.8 ± 18.8	43.8 ± 19.8
Vigorous (min·d^{-1})			
GPAQ	5.0 ± 15.5	22.2 ± 26.0	64.3 ± 43.0
IPAQ	21.9 ± 31.4	45.4 ± 48.5	55.0 ± 34.4
ActiGraph	3.0 ± 9.0	2.9 ± 5.1	3.0 ± 4.0
Moderate-Vigorous (min·d^{-1})			
GPAQ	44.4 ± 31.0	83.9 ± 38.9	187.2 ± 87.6
IPAQ	64.7 ± 51.4	102.7 ± 76.8	111.8 ± 68.0
ActiGraph	29.9 ± 22.1	37.7 ± 20.0	46.9 ± 20.4
Yamax			
Steps·d^{-1}	9225 ± 2425	11218 ± 2077	12363 ± 2035
Indirect Criterion Measures			
Anthropometric			
BMI (kg/m^2)	27.6 ± 4.0	27.1 ± 6.3	26.9 ± 8.3
Waist Circumference (cm)	93.0 ± 12.1	90.0 ± 14.9	85.4 ± 16.8
Body Fat (%)	36.2 ± 7.7	33.0 ± 9.3	28.4 ± 13.7
Fitness			
Estimated VO$_2$ max (ml·kg^{-1}·min^{-1})	27.8 ± 8.1	29.9 ± 8.1	36.2 ± 17.2
Steady State HR (b·min^{-1})	131.8 ± 7.8	134.9 ± 9.5	133.6 ± 8.8
Resting HR (mmHg)	77.0 ± 9.0	77.1 ± 11.9	74.9 ± 12.8
Systolic BP (mmHg)	117.0 ± 10.6	113.4 ± 11.7	117.6 ± 15.6
Diastolic BP (mmHg)	81.3 ± 9.5	77.2 ± 8.8	79.3 ± 11.8

Note. All differences by GPAQ category were not significant ($p > .05$)

Table 3 shows the correlations between the GPAQ min·d^{-1} and the validation variables. GPAQ scores were related with the accelerometer minutes of moderate ($r = 0.28$, p = 0.04), vigorous ($r = 0.48$, $p < 0.01$) and MVPA ($r = 0.26$, p = 0.04). The GPAQ and IPAQ were correlated for minutes of activity at each intensity category ($r = 0.04$ to 0.63). Minutes in MVPA were related to percent body fat ($r = -0.32$, $p < 0.01$), waist circumference ($r = 0.26$, p = 0.04), and estimated VO_2 max ($r = 0.26$ p = 0.04). Other correlations were low and not statistically significant ($p > 0.05$).

Table 3

Correlation Coefficients Between GPAQ Min·d^{-1} by Lntensity Level and Validation Measures: Study 1

	GPAQ Inactive	GPAQ Moderate	GPAQ Vigorous	GPAQ Moderate -to- Vigorous
Indirect Criterion Measures of Physical Activity				
Body Mass Index (wt kg/ht m^2)^a	.03	-.14	-.09	-.18
Body Fat^a	.02	-.15	-.25*	-.27*
Waist circumference (cm)^a	.01	-.16	-.20	-.26*
Estimated VO_2 max (ml·kg^{-1}·min^{-1})^a	.09	.16	.25*	.26*
Resting heart rate (b·min^{-1})^a	.02	.04	.01	-.00
Systolic blood pressure (mmHg)^a	-.02	-.04	-.11	-.11
Diastolic blood pressure (mmHg)^a	.03	-.12	-.13	-.16
Direct Criterion Measures of Physical Activity				
Pedometer (steps·d^{-1})^a	.03	.31*	.16	.39*
ActiGraph (min·d^{-1})$^{b\ c}$				
Inactive	-.12	-.20	-.13	-.29*

Light	.02	.09	-.12	.09
Moderate	.15	.28*	-.07	.27*
Vigorous	.18	-.18	.48*	.20
Moderate-Vigorous	.17	.20	.01	.26*
IPAQ (min·d^{-1})[a]				
Sitting	.51**	-.14	-.15	-.15
Walking				
Moderate	-.17	.04	.19	.13
Vigorous	.06	.04	.63**	.43**
MVPA[d]	-.18	.26**	.42**	.48**

[a] *n* = 69; [b] *n* = 53; * p < .05; ** p < .01 [c]ActiGraph Cut-points: Inactive (0-99 cts·min^{-1}), Light (101-1951 cts·min^{-1}) Moderate (1952-5724 cts·min^{-1}) Vigorous (5725+ cts·min^{-1}) Moderate-Vigorous (1952+ cts·min^{-1}); [b] sum of IPAQ walking, moderate, and vigorous

[d] Sum of IPAQ moderate and vigorous

Table 4 shows the percent agreement and kappa coefficients between the GPAQ and the IPAQ categories of low, moderate, and high physical activity levels. Agreement was higher in the higher intensity groups (low=25%, moderate=51.3%, and high=61.3%). The overall kappa value indicated slight-to-fair agreement between the activity categories.

Table 4

Categorical Agreement of the GPAQ and Short IPAQ: Study 1 (n=69)

	GPAQ n (%)	IPAQ n (%)	% Agreement	Kappa [a] (95% CI)
Low	14 (20.3%)	9 (13.0%)	25.0	
Moderate	39 (56.5%)	33 (47.8%)	51.3	.21 (.04-.39)
High	16 (23.2%)	27 (39.1%)	64.3	

[a] Cohen's weighted Kappa coefficient.

Bland-Altman plots between the GPAQ, IPAQ, and the ActiGraph are shown in Figure 1. Comparison of the GPAQ and IPAQ with the ActiGraph moderate- and vigorous-intensity physical activity showed increasing bias with higher levels of over reporting physical activity. A similar systematic bias was observed for both the GPAQ and IPAQ compared with the ActiGraph accelerometer (Figure 1b and 1c, respectively)

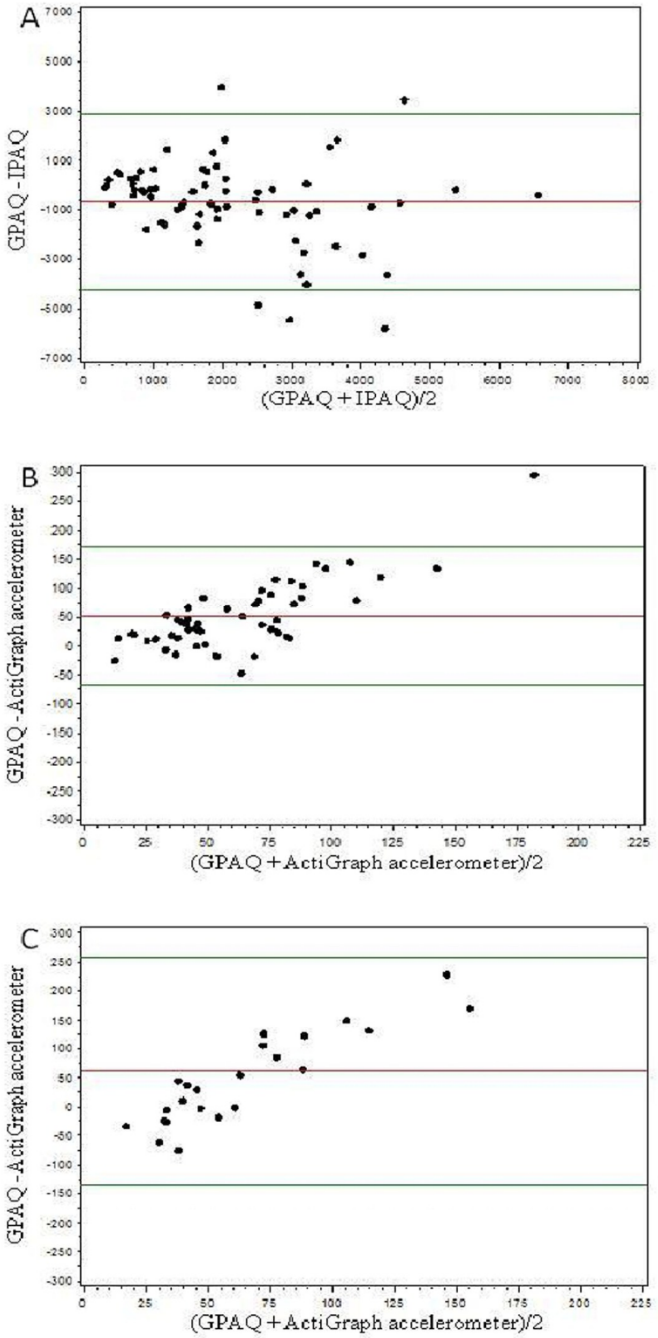


Figure 1. Bland Altman plot for: (a) the difference between the MET-min/week total scores for the short IPAQ and GPAQ: study 1 data; (b) moderate GPAQ minutes per day and the moderate ActiGraph minute per

day scores: study 1 data; (c) vigorous GPAQ minutes per day and the vigorous ActiGraph minute per day scores: study 1 data.

Study 2: Reliability

The 10-day test-retest reliability data produced good-to-excellent results for each activity level and activity categories ranging from *r* = 0.83 to 0.96 (Table 5).

Table 5

Mean ± SD and Interclass Correlation Coefficients for the 10-day Test-retest Reliability of the GPAQ by Intensity Levels and by Physical Activity Domains: Study 2 (n=16)

	Test 1	Test 2	ICC (95% CI)
GPAQ Activity Category ($min^{.}d^{-1}$)			
Moderate	92 ± 84	81 ± 63	.88 (.65-.96)
Vigorous	29 ± 31	25 ± 32	.84 (.53-.94)
Total	121 ± 98	106 ± 77	.89 (.68-.96)
GPAQ Domain ($min^{.}d^{-1}$)			
Travel	6 ± 15	9 ± 14	.83 (.49-.94)
Work: Moderate	44 ± 84	29 ± 54	.87 (.63-.96)
Work: Vigorous	2 ± 8	0 ± 0	n/a
Recreation: Moderate	42 ± 38	43 ± 41	.96 (.89-.99)
Recreation: Vigorous	27 ± 29	25 ± 32	.90 (.72-.97)
Sedentary	510 ± 179	568 ± 161	.92 (.78-.97)

Discussion

According to the World Health Organization, the goal of the GPAQ is to make intra- and inter-country comparisons of the prevalence of physical activity and sedentary behaviors and to compute estimates for the attributable risk of physical inactivity for chronic disease conditions (Bauman & Craig, 2005; WHO, 2002). Two versions of the GPAQ have been developed. Version 1 was evaluated by Armstrong and Bull (2006) in 2,657 adults from nine countries, and showed moderate criterion validity evidence for the total physical activity score compared with pedometers ($r = 0.31$), moderate-to-good convergent validity against the IPAQ ($r = 0.54$), and good-to-excellent 3- to 7-day test-retest reliability ($r = 0.67$ to 0.81) in a diverse sample of men and women (Armstrong & Bull, 2006; Bull et al., 2009). Following discussion by experts, Version 2 of the GPAQ was made slightly shorter than version 1 and was created for use in the WHO STEPS for risk factor surveillance among its member countries. This paper presents the results of two studies that describe the validity and reliability of Version 2 of the GPAQ in a sample of healthy adults.

In the current study, the GPAQ showed modest validity (r = 0.25 to 0.63) against measures of physical fitness (cardiorespiratory fitness, body composition), objective (accelerometer, pedometer), and subjective measures of physical activity (IPAQ). The highest correlations were seen between the accelerometer and GPAQ moderate, vigorous, and MVPA minutes and the GPAQ and IPAQ vigorous and MVPA minutes. Similarly, in an adult Vietnamese population during the wet and dry season, Trinh et al. (2009) demonstrated GPAQ correlations with an accelerometer ranging from .20 to .34 (25). Furthermore, our results are comparable to a report by Sallis et al. (2000) that showed validity evidence for seven self-reported questionnaires which had correlations ranging from 0.14 to 0.53 against objective criterion validity measure such as doubly labeled water, accelerometers (and other motion sensors), direct observation, and heart rate monitoring (Sallis & Saelens, 2000). Additionally, in the 12 country study using the short IPAQ, Craig et al. (2003) showed pooled validity correlations of about 0.30 (95% CI .23-.36) when compared to accelerometers. The IPAQ has since been used to assess PA in surveillance studies and to reflect PA behaviors in other study designs globally (Bauman, et al., 2009; Guthold, et al., 2008; Hagstromer, et al., 2006; Rutten, et al., 2003).

Test-retest reliability (see table 5) over 10 days was high (ICC = 0.83 to 0.96). These findings are slightly higher than previous studies investigating the test-retest reliability of the IPAQ (0.34 to 0.93) (Brown, Trost, Bauman, Mummery, & Owen, 2004; Craig, et al., 2003), GPAQ version 1 (0.67 to 0.81) (Armstrong & Bull, 2006; Bull, et al., 2009; Trinh et al., 2009). The GPAQ may show strong test-retest reliability because it differentiates activity according to domains (occupation, transportation, leisure). Respondents may provide more consistent answers because they can more easily separate types of activity. However, it is possible that this may allow for additional over reporting of time in activity due to added opportunities to report the type of physical activity performed.

The GPAQ showed an overestimation by approximately two to three times the amount of moderate activity that was recorded on the accelerometer and vastly over estimated the amount of vigorous activity recorded on the accelerometer. It is possible that some of this overestimation is due to the use of a waist mounted accelerometer underestimating non-ambulatory activities (Hendelman, et al., 2000) that could be assessed on the GPAQ while also generally over-reporting physical activity on the GPAQ. These concerns about overestimation are similar to previous studies that have identified an under reporting of light intensity physical activity and an over reporting of moderate and vigorous intensity physical activity (Bauman, et al., 2009; Boon, Hamlin, Steel, & Ross, 2010; Hagstromer, et al., 2006; Klesges, Eck, Mellon, Fulliton, Somes, & Hanson, 1990; Rutten, et al., 2003; Rzewnicki, Auweele, & Bourdeaudhuij, 2003; Sallis & Saelens, 2000).

Similar to the IPAQ, the GPAQ has categories to classify respondents' physical activity levels as low, moderate, and high. The GPAQ categories showed graded increases in VO_2 max, steps, accelerometer moderate and MVPA minutes, and IPAQ vigorous and MVPA minutes. A graded decrease was observed in percent body fat, waist circumference, and accelerometer sedentary minutes across GPAQ categories. While these categorical differences were not statistically significant ($p > 0.05$), when used as a large-scale surveillance instrument as it is intended (Armsrtong & Bull, 2006) these differences (observed in table 2) may be clinically relevant and achieve statistical significance. In the International Prevalence Study, the IPAQ was used to classify the physical activity levels of people in 20 countries (Bauman, et al., 2009). The IPAQ showed that approximately 62% of the United States population was highly active and only 16% were classified as low active (Bauman, et al., 2009). In the current study, the GPAQ classified fewer adults as highly active (23%) and more adults as low active (20%) than the IPAQ (high = 39.1% and low 13%) (see Table 4). The percent agreement between the GPAQ and IPAQ was highest in the moderate and highly active groups. This difference in categorical classification may reflect differences in the time frame the respondents are asked to recall (usual week vs. last 7 days) and the mode of administrations (IPAQ was self-administered by website; GPAQ was telephone administered). The IPAQ is a seven day recall while the GPAQ inquires about typical physical activity. However, Craig et al. (2003) reported that versions of the IPAQ showed comparable results when asking participants about the “last 7 days” and a “usual week”. The differences in characterizing physical activity levels remains a concern for comparing the results of research and surveillance studies using different questionnaires and warrants further investigation into the variation in these similar questionnaires.

A limitation of study one is that this sample was part of a worksite health intervention program that was designed to increase walking. Data were collected during the first month of this study and may be capturing people during a more active period. An additional limitation in study two is the different in administration methods of the IPAQ (Online) and GPAQ (Telephone) questionnaires. Future research is needed to assess the impact of administering the GPAQ by an interviewer or as a self-administrated version.

Summary

This study investigated the validity and reliability of the GPAQ version 2. The GPAQ showed strong evidence of test-retest reliability and modest validity evidence. These results are comparable to other subjective physical activity questionnaires that have been recommended for use in monitoring population physical activity levels in adults.

References

Armstrong, T. & Bull, F. (2006). Development of the World Health Organization Global Physical Activity Questionnaire (GPAQ). *Journal of Public Health,14*(2), 66-70.

Bauman, A., Bull, F., Chey, T., Craig, C., Ainsworth, B., Sallis, J., et al. (2009). The International Prevalence Study on Physical Activity:

results from 20 countries. *International Journal of Behavioral Nutrition and Physical Activity, 6*(1), 21.

Bauman, A. & Craig, C. (2005). The place of physical activity in the WHO Global Strategy on Diet and Physical Activity. *International Journal of Behavioral Nutrition and Physical Activity, 2*(1), 10.

Bauman, A. E. & Sallis, J. F. (2008). Global problems require global studies. *American Journal of Preventive Medicine, 34*(6), 544-545.

Boon, R. M., Hamlin, M. J., Steel, G. D. & Ross, J. J. (2010). Validation of the New Zealand Physical Activity Questionnaire (NZPAQ-LF) and the International Physical Activity Questionnaire (IPAQ-LF) with accelerometry. *British Journal of Sports Medicine, 44*(10), 741-746.

Brown, W. J., Trost, S. G., Bauman, A., Mummery, K., & Owen, N. (2004). Test-retest reliability of four physical activity measures used in population surveys. *Journal of Science and Medicine in Sport*, *7*(2), 205-215.

Bull, F. C, Maslin, T. S. & Armstrong, T. (2009). Global physical activity questionnaire (GPAQ): nine country reliability and validity study. *Journal of Physical Activity and Health, 6*(6), 790-804.

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*(1), 155–159.

Craig, C. L., Marshall, A. L., Sjöström, M., Bauman, A. E., Booth, M. L., Ainsworth, B. E., et al. (2003). International physical activity questionnaire: 12-country reliability and validity. *Medicine and Science in Sports and Exercsie, 35*(8), 1381-1395.

Crouter, S. E., Schneider, P. L., Karabulut, M. & Bassett, D. R. (2003). Validity of 10 electronic pedometers for measuring steps, distance, and energy cost. *Medicine and Science in Sports and Exercsie, 35* (8), 1455-60.

Freedson, P. S., Melanson, E., & Sirard, J. (1998). Calibration of the Computer Science and Applications, Inc. accelerometer. *Medicine and Science in Sports and Exercise*, *30*(5), 777-781.

Guthold, R., Ono, T., Strong, K. L., Chatterji, S., & Morabia, A. (2008). Worldwide Variability in Physical Inactivity: A 51-Country Survey. *American Journal of Preventive Medicine, 34*(6):486-494.

Hagströmer, M., Oja, P., & Sjöström, M. (2006). The International Physical Activity Questionnaire (IPAQ): a study of concurrent and construct validity. *Public Health Nutrition*, *9*(6), 755-762.

Hendelman, D., Miller, K., Baggett, C., Debold, E., & Freedson, P. (2000) Validity of accelerometry for the assessment of moderate intensity physical activity in the field. *Medicine and Science in Sports and Exercise, 32*, 442–449.

Herrmann, S. D., Hart, T. L., Lee, C. D., & Ainsworth, B. E. (2011). Evaluation of the MyWellness Key accelerometer. *British Journal of Sports Medicine, 45*:109-113.

Jacobs, D. R., Ainsworth, B. E., Hartman, T. J., & Leon, A. S. (1993). A simultaneous evaluation of 10 commonly used physical activity questionnaires. *Medicine and Science in Sports and Exercise*, *25*(1), 81-91.

Klesges, R. C., Eck, L. H., Mellon, M. W., Fulliton, W., Somes, G. W., & Hanson, C. L. (1990). The accuracy of self-reports of physical activity. *Medicine and Science in Sports and Exercise, 22*(5), 690-697.

Matthews, C. E. (2005). Calibration of accelerometer output for adults. *Medicine and Science in Sports and Exercise, 37*, S512-522.

Matthews, C. E., Ainsworth, B. E., Hanby, C., Pate, R. R., Addy, C., Freedson, P. S., Jones, D. A., et al. (2005). Development and testing of a short physical activity recall questionnaire. *Medicine and Science in Sports and Exercise*, *37*(6), 986-994.

Rütten, A., Ziemainz, H., Schena, F., Stahl, T., Stiggelbout, M., Vanden Auweele, Y., et al. (2003). Using different physical activity measurements in eight European countries. Results of the European Physical Activity Surveillance System (EUPASS) time series survey. *Public Health Nutrition, 6*(4), 371-376.

Rzewnicki, R., Vanden Auweele, Y., & De Bourdeaudhuij, I. (2003). Addressing overreporting on the International Physical Activity Questionnaire (IPAQ) telephone survey with a population sample. *Public Health Nutrition*, *6*(3), 299-305.

Sallis, J. F. & Saelens, B. E. (2000). Assessment of physical activity by self-report: status, limitations, and future directions. *Research Quarterly for Exercise & Sport, 71*(2):S1-14.

Schneider, P. L., Crouter, S. E., & Bassett, D. R. (2004). Pedometer measures of free-living physical activity: comparison of 13 models. *Medicine & Science in Sports & Exercise*, *36*(2), 331-335.

Thomas, S., Reading, J., & Shephard, R. J. (1992). Revision of the physical activity readiness questionnaire (PAR-Q). *Canadian Journal of Sports Science, 17*(4)*,* 338-345.

Trinh, O. T. H., Nguyen, N. D., van der Ploeg, H. P., Dibley, M. J., & Bauman, A. (2009). Test-retest repeatability and relative validity of the Global Physical Activity Questionnaire in a developing country context. *Journal of Physical Activity and Health*, *6(*1), S46-53.

Troiano, R. P., Berrigan, D., Dodd, K. W., Mâsse, L. C., Tilert, T., & Mcdowell, M. (2007). Physical Activity in the United States Measured by Accelerometer. *Medicine & Science in Sports & Exercise*, *40*(1):181-188.

Tudor-Locke, C., McClain, J. J., Hart, T. L., Sisson, S. B., & Washington, T. L. (2009). Pedometry Methods for Assessing Free-living Youth. *Research Quarterly for Exercise & Sport, 80*(2), 175-184.

World Health Organization. (2002). The World Health Report 2002: Reducing risks, promoting healthy lifestyles. Geneva, Switzerland: World Health Organization. 248.

Chapter 6

HOW MANY HOURS ARE ENOUGH? OPTIMAL ACCELEROMETER WEAR TIME TO REFLECT DAILY ACTIVITY

Abstract

There is little consensus on how many hours of accelerometer wear time is needed to reflect a usual day. This study identifies the optimal $h \cdot d^{-1}$ of accelerometer wear time to assess daily physical activity. 124 adults (age = 41 ± 11 years; BMI = 27 ± 7 $kg \cdot m^{-2}$) contributed approximately 1,200 days accelerometer wear time. Five 40 day samples were randomly selected with 10, 11, 12, 13, and 14 $h \cdot d^{-1}$ of wear time. Four semi-simulation data sets (10, 11, 12, 13 $h \cdot d^{-1}$) were created from the criterion 14 $h \cdot d^{-1}$ data set to assess Absolute Percent Error (APE). Repeated measures ANOVAs compared $min \cdot d^{-1}$ between 10, 11, 12, 13 $h \cdot d^{-1}$ and the criterion 14 $h \cdot d^{-1}$ for inactivity (<100 $cts \cdot min^{-1}$), light (100–1951 $cts \cdot min^{-1}$), moderate (1952–5724 $cts \cdot min^{-1}$), and vigorous (≥5725 $cts \cdot min^{-1}$) physical activity. APE ranged from 5.6 to 41.6% (10 $h \cdot d^{-1}$ = 28.2-41.6%; 11 $h \cdot d^{-1}$ = 20.3-36.0%; 12 $h \cdot d^{-1}$ = 13.5-14.3%; 13 $h \cdot d^{-1}$ = 5.6-7.8%). $Min \cdot d^{-1}$ differences were observed for inactivity, light, and moderate physical activity between 10, 11, 12, and 13 $h \cdot d^{-1}$ and the criterion ($p < .05$). This suggests a minimum accelerometer wear time of 13 $h \cdot d^{-1}$ is needed to provide a valid measure of daily physical activity.

Introduction

Adequate amounts of physical activity have been shown to offer wide-ranging health benefits and to reduce the negative consequences of physical inactivity (Warburton, et al., 2006). Recently, the United States Department of Health and Human Services provided guidelines for the amount and type of physical activity that offers health benefits (USDHHS, 2008). In order to understand whether people are meeting these guidelines, it is imperative to use highly accurate assessment methods, such as accelerometers, to measure physical activity.

Accelerometers are objective physical activity monitors that directly measure the duration, intensity, and frequency of physical activity (Welk, et al., 2000). Physical activity research using accelerometers has become more common with approximately 1,200 articles cited in the PubMed database since 2000. This wealth of knowledge has provided general consensus on certain aspects of accelerometer-measured physical activity. Recommendations exist for the number of days that are needed to monitor free-living physical activity (Trost et al., 2005; Ward, Evenson, Vaughn, Rodgers, & Troiano, 2005) and optimal epoch lengths to capture activity in adults and children (McClain, Abraham, Brusseau, Tudor-Locke, 2008; Trost, et al., 2005). However, current research provides little evidence on how many hours of accelerometer wear time is needed each day to reflect a day’s activity in an analysis (Murphy 2009).

Recent studies have reported a variety of criteria for selecting the number of hours per day one should wear an accelerometer to reflect daily free-living physical activity (Catellier, et al., 2005; Frank et alk., 2005; Matthews et al., 2008; Slootmaker, et al., 2009; Troiano, et al., 2007; Young et al., 2009). These studies have hours per day values ranging from as few as 6 $h \cdot d^{-1}$ (Young, et al., 2009) and up to 16 $h \cdot d^{-1}$ (Slootmaker, et al., 2009) to constitute a valid day. Slootmaker et al. (2009) used an assumption that people sleep 8 $h \cdot d^{-1}$ and therefore restricted valid days to persons with less than 16 hours of monitoring. Results from the 2004 accelerometer consensus meeting suggest using the 70/80 rule (Catellier, et al., 2005) for required daily wear time (Ward, et al., 2005). This rule provides a sample specific recommendation based on 70% of the sample having accelerometer data with at least 80% of those having at least the same amount of hours. Another approach that has been used to determine accelerometer hours per day wear time is to normalize each person's total minutes of daily activity to a 12 $h \cdot d^{-1}$ to balance different amounts of wear time (Young, et al., 2009). This approach may create an over- or underestimation of actual movement time. A common approach, including that used in U.S. National Health and Nutrition Examination Survey (NHANES) accelerometer analyses, is to require 10 or more $h \cdot d^{-1}$ of accelerometer wear time to be considered a valid measurement day (Matthews, et al., 2008; Troiano, et al., 2007).

To date, empirical evidence has not been published that recommends a minimal number of hours per day needed to reflect optimal daily wear time for accelerometers. This is important because wearing an accelerometer for too few hours per day can result in an underestimation of time spent in different physical activity intensity categories. Accordingly, requiring study participants to wear an accelerometer longer than is needed to obtain sufficient information to reflect usual physical activity patterns can cause undue burden on study participants. In this paper, a semi-simulation approach was utilized to create missing data to identify the Absolute Percent Error (APE) between varying hours per day accelerometers were worn by adults enrolled in a worksite physical activity study. APE is a useful method to compare actual wear time data (criterion) and comparison data (semi-simulated). The purpose of the analyses was to identify the optimal number of hours per day an accelerometer should be worn to obtain a valid measure of daily free-living physical activity.

Methods

Study Design and Participant Selection

Study participants (*n* = 152) were randomly selected to complete periodic physical fitness tests from 714 participants enrolled in a six-month, quasi-experimental (lack of a control group) worksite health promotion study designed to increase daily walking. The goal of the walking study was to increase physical activity by taking 10,000 steps per day as recorded by a pedometer. During the first month of the intervention study, 124 participants from the fitness test group wore an accelerometer for at least one week to identify their frequency and duration of physical activity at varying intensity levels. The remaining 28 participants from the fitness group lacked accelerometer data and were excluded from data analyses for this study. Baseline data from participants who wore an accelerometer (age = 41 ± 11 yrs; BMI = 27 ± 7 $kg \cdot m^{-2}$) were used for the data analyses. There were no differences in age or BMI between those who wore the accelerometers and those without accelerometer data ($p > .05$). Collectively, the 124 study participants wearing the accelerometers contributed approximately 1,200 days of accelerometer monitoring data.

Prior to completing any study activities, participants read and signed an informed consent form approved by the university Institutional Review Board. Exclusion criteria from the worksite health promotion study were determined using the PAR-Q (Thomas, et al., 1992) and included the inability to walk, current diagnosis of cardiovascular disease, uncontrolled risk factors for cardiovascular disease (e.g., uncontrolled hypertension), signs and symptoms of cardiovascular intolerance to exercise (e.g., dizziness with exertion), and current pregnancy. Of those volunteering for the worksite health promotion study, none were excluded based on the exclusion criteria. Table 1 shows the characteristics of the accelerometer monitoring sub-sample (*n* = 124).

Table 1

Means and Standard Deviations of Descriptive Variables for the Study Participants

	Total (*n* = 124)	Women (*n* =96)	Men (*n* =28)
Age	41 ± 11	42 ± 11	39 ± 13
BMI	27.4 ± 7.1	26.9 ± 6.2	30.3 ± 9.9

Physical fitness and accelerometer data were obtained during a scheduled laboratory visit whereby participants completed tests to obtain (in this order) age, height, weight, percent body fat, waist circumference, resting heart rate, resting blood pressure, and estimated VO_2 max. At the end of the testing session, participants were given an ActiGraph GT1M accelerometer to wear for the next seven days. Data from the physical fitness measures were not used for the current paper and are not discussed.

Accelerometry

The ActiGraph model GT1M accelerometer (ActiGraph, LLC, Pensacola, Florida, USA) is a uniaxial piezoelectric accelerometer (3.8cm x 3.7cm x 1.8cm; 27 grams) that assesses physical activity duration at varying intensities of movement expressed as numeric counts (higher counts reflect higher intensity of movement with zero denoting no movement), steps, and an estimate of caloric expenditure. The GT1M records vertical accelerations ranging in magnitude from approximately 0.05 to 2.0G with a frequency response from 0.25 to 2.50 Hz. Output data are digitized at a rate of thirty times per second with intensity data recorded in one minute epochs (sampling interval). Step counts and the caloric estimates were not recorded for the present study. Data were downloaded using ActiGraph software and stored in a computer database (ActiGraph 2009).

Upon receiving the ActiGraph, participants were instructed to wear the accelerometer for the next seven days upon waking (or immediately after bathing or showering) until they retired at night. The recording was activated to run from midnight the night subjects received the accelerometer until midnight seven days later. The accelerometer was worn on the right hip attached to an elastic belt. The participants were asked to remove the accelerometer during showering/bathing and any water activities.

Data Management

Non-wear periods were identified as 60 consecutive minutes with no movement data (zero counts) allowing up to 2 consecutive minutes of 1-100 cts·min^{-1}(Matthews 2005). Non-wear periods were ended with >100 cts·min^{-1} or with 3 consecutive 1–100 cts·min^{-1} (Matthews 2005). ActiGraph monitors were scored to assess time spent in each intensity level using a SAS statistical program. Matthew’s cut-points were used to identify time spent in inactivity (0 – 99 cts·min^{-1}), light intensity (100-1951 cts·min^{-1}) (Matthews 2005) and Freedson's cut points were used to determine time spent in moderate intensity (1952-5724 cts·min^{-1}) and vigorous intensity (5725+ cts·min^{-1}) (Freedson, et al., 1998).

To determine the criterion wear time, recent accelerometer research studies using large numbers of participants were examined. An average accelerometer wear time of 14 h·d^{-1} was identified and used as the criterion value for a day's recording (criterion day) (Hagstromer, et al., 2007; Troiano, et al., 2007). The data were split into two samples (original and validation sample). The original sample of 40 days where 40 different participants wore the accelerometer for 14 valid hours (i.e., the accelerometer was worn for at least 40 min for each hour) (Evenson & Terry, 2009) was selected from the larger data set. This was used as reference for comparison of ActiGraph data using shorter hour per day specifications. Additional forty-day samples of 13, 12, 11, and 10 h·d^{-1} of accelerometer wear time were then selected and used as reference for missing data in the semi-simulation approach. The semi-simulation approach removed data from the known 14 hour criterion day data set by matching the missing data pattern of the 13, 12, 11, and 10 hour data sets. This was done to allow a comparison of the min·d^{-1} spent in activity at varying intensities by different hours of wear time. The data management procedure to create semi-simulated data sets is shown in figure 1.

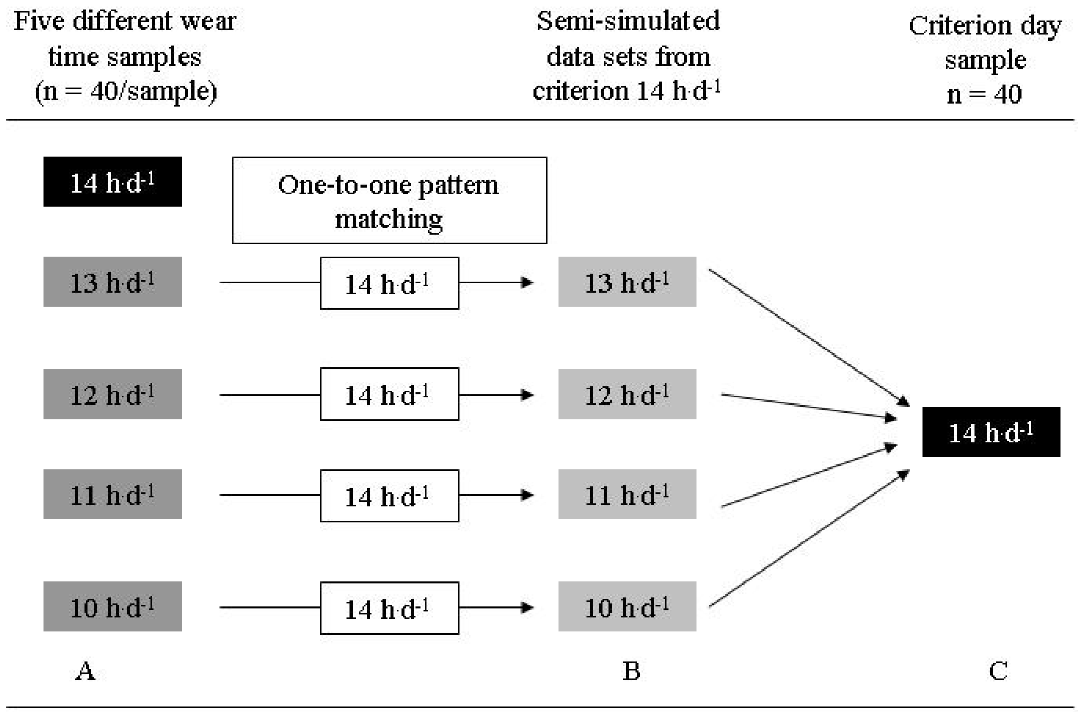


Figure 1. A = h·d^{-1} sample serves as a criterion; A-B = Data removed from 14 h·d^{-1} to match pattern of less wear time; and B-C = Compare data of semi-simulated data sets to criterion sample data.

This method was preferred because it uses data patterns of real subjects to remove data instead of a purely random data removal method. For example, a 14 h·d^{-1} recorded accelerometer counts from 6 a.m. to 8 p.m. and a 13 h·d^{-1} recorded counts from 7 a.m. to 8 p.m. To simulate a 13 h·d^{-1} pattern, accelerometer counts between 6 a.m. and 7 a.m. would be removed from the 14 h·d^{-1}. Similarly, for a 10 h·d^{-1}, if an individual wore the accelerometer between 9 a.m. and 8 p.m., but also had missing data from between 1 p.m. and 2 p.m. (10 hour total wear time) then data would be removed from the 14-hour day between 6 a.m. and 9 a.m. and between 1 p.m. and 2 p.m. Figure 2 depicts the semi-simulation approach for a 13 h·d^{-1} example.

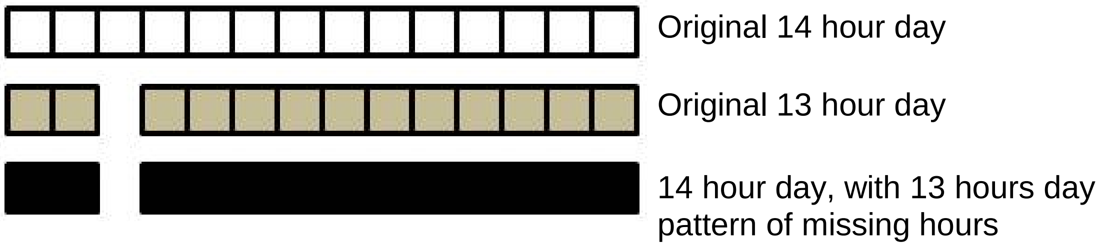


Figure 2. Example to create a semi-simulation data set of13 h·d^{-1}. A = Original 14 h·d^{-1} pattern (criterion day); B = Original 13 h·d^{-1} pattern; C = 14 h·d^{-1}, with 13 h·d^{-1} pattern of missing hours.

This procedure was repeated to create data sets of 12, 11, and 10 hours. In addition, a validation sample was created by repeating the semi-simulation process on a separate random sample of individuals and days. This was done to compare the accuracy of the modeling performed with the original accelerometer data sample. More information about the semi-simulation approach can be found elsewhere (Kang, et al., 2009; Kang, et al., 2005).

Data Analysis

The semi-simulated approach was used to compare the criterion day (14 h·d^{-1}) data set to semi-simulation data sets of 13, 12, 11, and 10 h·d^{-1} of accelerometer wear time. Accordingly, repeated measures ANOVAs with the Least Significant Difference (LSD) post hoc method were performed to assess differences in daily minutes between semi-simulation data sets (13, 12, 11, and 10 h·d^{-1}) and criterion day (14 h·d^{-1}) at various intensity levels: inactivity (<100 cts·min^{-1}), light (100–1951 cts·min^{-1}), moderate (1952–5724 cts·min^{-1}), and vigorous (≥5725 cts·min^{-1}). APE was computed between the criterion day and each of the semi-simulation data sets for daily minutes of inactivity (<100 cts·min^{-1}), light (100–1951 cts·min^{-1}), moderate (1952–5724 cts·min^{-1}), and vigorous (≥5725 cts·min^{-1}) activity. APE identifies the percent difference between two values with a lower APE desired (Kang, et al., 2009). Repeated measures ANOVAs and APE also were computed for the validation sample to cross validate the results of the original sample. The alpha level was set at .05. Microsoft Excel 2007 and SPSS 17.0 (SPSS Inc., Chicago, Illinois) were used for the data analysis.

Results

Table 2 presents the average min·d^{-1} of activity for each accelerometer wear time duration (10, 11, 12, 13, and 14 h·d^{-1}) obtained from the original sample and from the separate validation sample. In general, the minutes spent in each intensity level increased with increasing hours of accelerometer wear time. This finding was consistent in the original and in the validation sample. Repeated measures ANOVA results showed significant differences in daily minutes between 10, 11, 12, and 13 $h \cdot d^{-1}$ and the criterion 14 $h \cdot d^{-1}$ for inactivity, light intensity, and moderate intensity activity (all $p < .05$). No difference was found for vigorous intensity activity, F (2.34, 91.09) = 2.06, p = .13. Similar results were found in the validation sample with no significance difference in daily minutes for the vigorous intensity activity, F (2.48, 96.71) = 1.7, p = .17, and significant differences in all other intensity levels (all $p < .05$). The greatest difference in $min \cdot d^{-1}$ for each intensity level was seen between the criterion day (14 $h \cdot d^{-1}$) and 10 $h \cdot d^{-1}$ recording time. There was not a significant difference between the original and validation sample in any category (all $p > .05$).

Table 2

Mean and Standard Deviation for the Minutes of Physical Activity by Intensity and Hours per Day of Accelerometer Wear Time

Hours of data	Activity Classification	Original Sample		Validation Sample[a]	
		Mean	SD	Mean	SD
14 hours	Inactivity	567.3	67.7	578.1	101.5
	Light Activity	212.9	63.7	218.7	97.0
	Moderate Activity	45.5	26.1	83.7	28.2
	Vigorous Activity	5.2	11.8	5.3	9.2
13 hours[b]	Inactivity	527.6	63.8	537.0	93.9
	Light Activity	196.0	58.8	203.9	89.5
	Moderate Activity	43.0	26.1	77.7	26.8
	Vigorous Activity	5.0	11.8	4.6	8.0

12 Hours[b]	Inactivity	486.6	60.6	503.3	81.1
	Light Activity	183.0	56.7	189.6	82.4
	Moderate Activity	39.1	25.3	71.4	24.7
	Vigorous Activity	4.0	7.4	4.4	7.9
11 Hours[b]	Inactivity	451.4	52.6	461.8	79.5
	Light Activity	164.7	52.7	173.7	77.4
	Moderate Activity	34.4	24.0	66.5	24.2
	Vigorous Activity	3.5	10.4	4.1	8.7
10 Hours[b]	Inactivity	406.5	47.8	415.9	72.5
	Light Activity	151.2	43.0	161.5	71.7
	Moderate Activity	33.3	25.4	60.9	23.2
	Vigorous Activity	3.1	8.1	3.6	7.4

Notes. [a]No difference between Original and Validation sample (all $p > .05$). [b]All activity intensities were significantly different from the 14 hour criterion ($p < .05$) except vigorous intensity.

The APE values for the accelerometer wear time data are presented in Table 3. APE values were similar between the original sample and the validation sample. The range of APE values increased with the shorter accelerometer recording time in the original (13 h·d^{-1} = 5.6 to 7.8%; 12 h·d^{-1} = 13.5 to 14.3%, 11 h·d^{-1} = 20.3% to 36.0%; 10 h·d^{-1} = 28.2% to 41.7%) and validation sample (13 h·d^{-1} = 7.0 to 8.5%; 12 h·d^{-1} = 13.7 to 16.5%, 11 h·d^{-1} = 20.2% to 26.9%; 10 h·d^{-1} = 25.6% to 36.2%). A wear time of 13 h·d^{-1} was the only category with APE for each intensity level less than 10%.

Table 3

Absolute Percentage Error (APE) for the Original and Validation Samples by Intensity and Hours per Day of Accelerometer Wear Time

Hours of data	Activity Classification	Original Sample		Validation Sample	
		Mean	SD	Mean	SD
13 hours	Inactivity	7.02%	2.13%	7.04%	2.08%
	Light Activity	7.84%	4.74%	6.66%	4.79%
	Moderate Activity	6.36%	11.02%	6.64%	7.93%
	Vigorous Activity	5.63%	13.23%	8.48%	20.68%
12 Hours	Inactivity	14.21%	3.00%	14.41%	3.93%
	Light Activity	14.16%	5.53%	14.65%	7.90%
	Moderate Activity	14.34%	22.31%	13.74%	12.31%
	Vigorous Activity	13.49%	24.96%	16.52%	31.84%
11 Hours	Inactivity	20.28%	3.73%	20.53%	4.64%
	Light Activity	22.82%	6.88%	20.95%	5.29%
	Moderate Activity	24.33%	26.30%	20.16%	11.88%
	Vigorous Activity	36.01%	43.48%	26.92%	35.50%
10 Hours	Inactivity	28.20%	3.72%	27.90%	5.68%
	Light Activity	28.48%	7.78%	25.61%	9.43%
	Moderate Activity	29.19%	25.63%	27.21%	14.74%
	Vigorous Activity	41.66%	41.53%	36.20%	43.20%

Discussion

This study provides a scientific rationale for determining the optimal number of hours per day an accelerometer should be worn to obtain a valid measure of daily free-living physical activity. The results showed that allowing different amounts of accelerometer wear time had a significant impact on the amount of time assessed in various intensity levels and that wearing an accelerometer at least 13 $h \cdot d^{-1}$ had the lowest APE when compared to a criterion wear time of 14 $h \cdot d^{-1}$. For all intensity levels, except vigorous intensity, increases in the accelerometer wear time resulted in increasing amounts of activity per day. Time spent in inactivity, light, and moderate intensity activity was nearly 30% less when 10 $h \cdot d^{-1}$, the most commonly used criteria, was compared to the criterion 14 $h \cdot d^{-1}$ of accelerometer wear time. This demonstrates a significant underestimation of free-living physical activity and provides some insight as to why accelerometers have been often questioned for underestimation in certain instances (Crouter et al., 2006; Ham et al., 2007; Welk et al., 2000). These data show that using daily accelerometer data with less than 13 hours results in significantly reduced minutes of activity and higher than recommended error. These findings were supported in the validation sample which used separate randomly selected days.

The recent findings from NHANES found that adult Americans spent between 16-38 minutes (using all activity minutes) doing moderate intensity activity (Troiano et al., 2007). NHANES required at least 10 hours of wear time to be included in the analysis and had a mean wear time of 14.2 hours. Despite a mean wear time of 14 $h \cdot d^{-1}$, it is possible that the mean minutes of moderate intensity activity were biased toward underestimation by including data with 10, 11, and 12 $h \cdot d^{-1}$ of wear time. In 2008, Matthews et al. analyzed the 2003-04 NHANES accelerometer data using the same criteria as Troiano et al. (2007) with a minimum wear time of 10 h·d^{-1}. They reported a mean wear time 13.9 h·d^{-1} with inactivity durations ranging from about 430 to 570 min·d^{-1} (Matthews et al., 2008; Troiano et al., 2007). The study also observed a positive association between wear time and time in inactivity and therefore adjusted for mean wear time in the analysis. Interestingly, while the current study also observed an increase in inactive time ranging from 406.5 to 578.1 min·d^{-1} with increasing wear time, the proportion of time in inactivity remained unchanged despite the wear time requirements (~68%).

Min·d^{-1}of vigorous intensity physical activity was not statistically different between the hours per day of accelerometer wear times. While the mean minutes nearly doubled (3.1 to 5.2 min·d^{-1}) with an increased duration of wear time, the standard deviation was high due to very few participants performing vigorous intensity physical activity. Troiano et al. (2007) reported similar low levels of vigorous intensity activity in the NHANES study (0.1 to 3 min·d^{-1}) suggesting a 10 h·d^{-1} wear time may be sufficient to capture vigorous intensity activity in the general population. However, different results may arise in population groups who perform higher levels of vigorous intensity physical activity.

The differences between the hours per day of accelerometer wear time observed in the current study sample have significance for surveillance studies designed to assess the proportion of adults meeting physical activity guidelines, in intervention studies designed to assess changes in activity over time, and in studies designed to assess time spent in sedentary behaviors. Assessing physical activity for more hours per day can provide a more accurate picture of the true levels of physical activity and avoid misclassification bias. Misclassification of physical activity can bias results of studies toward the null, alter the interpretations for dose-response relationships between physical activity and inactivity exposures and health outcomes, modify the proportion of adults meeting public health physical activity recommendations, and change the interpretations of intervention studies (Aschengrau & Seage, 2003). As such it is important to standardize the accelerometer wear time required for study participants in various research and surveillance settings.

Some researchers have dealt with disparate wear time data by normalizing it to a standard time, such as 12 hours (Young, et al., 2009). Previously, there has been little justification for normalizing daily wear time and minutes of activity if an individual has less than a desired amount of wear time. Furthermore, not much is known about the accuracy of this method of data replacement and there is almost no evidence to support 12 hours as the standard time in this scenario. However, the data from this study show that the proportion of time spent in each activity category remains relatively stable which may support this method pending further study.

It is possible that requiring 13 hours of wear time is not always feasible. If so, then further efforts are needed to improve the daily assessment of physical activity. One option is to place a greater emphasis educating the participants on the importance of wearing the device during *all* waking hours. Sharpe et al (in press) showed it is possible to have high compliance in an accelerometer study by providing detailed instructions to study participants. Another alternative is to use the accelerometer for 24 hour surveillance. The area of sleep medicine and sleep research has employed accelerometry with great success (Sadeh & Acebo, 2002). Wearing accelerometers for 24 hour periods may provide richer data and help explain more complex relationships between sleep quality/quantity and physical activity and provide us with a more accurate measure of minutes of daily activity.

While this study provides data to show that a minimum accelerometer wear time of 13 $h \cdot d^{-1}$ is needed for an accurate measure of daily activity, this study has limitations which may restrict the external validity of the findings. First, the sample population (N = 714) was from faculty and staff of a large university (about 12,500 eligible employees) who were engaged in a worksite intervention study. Study participants included faculty, housekeeping staff, police officers, administrators, support staff, and others with varied work responsibilities. Despite this variability, this inclusion of participants from one university may have limited the variability of occupational physical activity behaviors. These findings may differ from those in a more diverse occupational sample. Second, enrollment in a pedometer-based walking study with a goal to accumulate 10,000 steps per day may have resulted in a study sample that was more active than the general population (Troiano, et al., 2007; Tudor-Locke, et al., 2009). The study sample took an average of 10,767 ± 3,265 steps per day which is more than the average 9,676 uncensored steps per day (6,540 censored steps per day) reported by Tudor-Locke et al. from the 2003-2004 NHANES data (Tudor-Locke, et al., 2009). Third, the mean age of participants in this study was 41 ± 11 yrs. It is unknown if the findings would differ in younger or older population age groups. Despite these limitations, this study is important because it is one of the first to specifically evaluate the amount of daily accelerometer wear time recommended to obtain a valid measure of physical activity. Furthermore, the analytical method for the current study was based on an evidence-based approach showing significantly more time was spent in inactive behaviors and in light and moderate intensity physical activity as the accelerometer wear time increased. As such, wearing an accelerometer at least 13 $h \cdot d^{-1}$ has the potential to impact recommendations for time spent in inactivity or other intensities of physical activity.

Summary

These data provide support for requiring at least 13 hours of wear time when analyzing accelerometer data when 14 $h \cdot d^{-1}$ was considered as the criterion day. The results also provide information about the amount of underestimation and error that can occur by allowing fewer hours per day of accelerometer data. It is not unexpected that wear time closer to the criterion is more accurate; however, it is clear that allowing less wear time in an analysis can significantly influence the results. This study has the potential to provide a standard recommendation for the number of hours per day that participants must wear accelerometers in physical activity research studies and in surveillance settings.

References

ActiGraph. (2009). *ActiLife Users Manual*. Pensacola, FL: ActiGraph; p. 8-51.

Aschengrau, A. & Seage, G. R. (2003). *Essentials of Epidemiology in Public Health Practice.* Subbury, MA: Jones and Bartlett Publishers; p. 251-279.

Catellier, D. J., Hannan, P. J., Murray, D. M., Addy, C. L., Conway, T. L., Yang, S., & Rice, J. C. (2005). Imputation of missing data when measuring physical activity by accelerometry. *Medicine and Science in Sports and Exercise, 37*, S555-S562.

Crouter, S. E., Churilla, J. R., & Bassett, D. R. (2006). Estimating energy expenditure using accelerometers. *European Journal of Applied Physiology, 98*, 601-612.

Evenson, K. R. & Terry, J. W. (2009). Assessment of differing definitions of accelerometer nonwear time. *Research Quarterly for Exercise and Sport*, *80*(2), 355-362.

Frank, L. D., Schmid, T. L., Sallis, J. F., Chapman, J., & Saelens, B. E. (2005). Linking objectively measured physical activity with objectively measured urban form: Findings from SMARTRAQ. *Am J Preventive Medicine, 28*(2), 117-125.

Freedson, P. S., Melanson, E., & Sirard, J. (1998). Calibration of the Computer Science and Applications, Inc. accelerometer. *Medicine and Science in Sports and Exercise*, *30*(5), 777-781.

Hagströmer, M., Oja, P., & Sjöström, M. (2007). Physical activity and inactivity in an adult population assessed by accelerometry. *Medicine and Science in Sports and Exercise, 39*(9), 1502-1508.

Ham, S. A., Reis, J. P., Strath, S. J., Dubose, K. D., & Ainsworth, B. E. (2007). Discrepancies between methods of identifying objectively determined physical activity. *Medicine & Science in Sports & Exercise, 39*, 52-58.

Kang, M., Bassett, D. R., Barreira, T. V., Tudor-Locke, C., Ainsworth, B. E., Reis, J. P., Strath, S., et al. (2009). How Many Days Are Enough? A Study of 365 Days of Pedometer Monitoring. *Research Quarterly for Exercise and Sport*, *80*(3), 445-453.

Kang, M., Rowe, D. A., Barreira, T. V., Robinson, T. S., & Mahar, M. T. (2009). Individual information-centered approach for handling physical activity missing data. *Research Quarterly for Exercise and Sport*, *80*(2), 131–137.

Kang, M., Zhu, W., Tudor-Locke, C., & Ainsworth, B. (2005). Experimental determination of effectiveness of an individual information-centered approach in recovering step-count missing data. *Measurement in Physical Education and Exercise Science, 9*, 233–250.

Matthews, C. E. (2005). Calibration of accelerometer output for adults. *Medicine and Science in Sports and Exercise, 37*, S512-522.

Matthews, C. E., Chen, K. Y., Freedson, Patty S., Buchowski, M. S., Beech, B. M., Pate, R. R., & Troiano, R. P. (2008). Amount of Time Spent in Sedentary Behaviors in the United States, 2003–2004. *American Journal of Epidemiology*, *167*(7), 875 -881.

McClain, J., Abraham, T., Brusseau, T., & Tudor-Locke, C. (2008). Epoch Length and Accelerometer Outputs in Children: Comparison to Direct Observation. *Medicine and Science in Sports and Exercise, 40*(12), 2080-2087.

Murphy, S. L. (2009). Review of physical activity measurement using accelerometers in older adults: Considerations for research design and conduct. *Preventive Medicine, 48*, 108–114.

Sadeh, A. & Acebo, C. (2002). The role of actigraphy in sleep medicine. *Sleep Medicine Reviews, 6*(2), 113-124.

Sharpe, P. A., Wilcox, S., Rooney, L. J., Strong, D., Hopkins-Campbell, R., & Butel, J. (In Press). Adherence to accelerometer protocols among women from economically disadvantaged neighborhoods. *Journal of Physical Activity and Health*.

Slootmaker, S. M., Schuit, A. J., Chinapaw, M. J., Seidell, J. C., & van Mechelen, W. (2009). Disagreement in physical activity assessed by accelerometer and self-report in subgroups of age, gender, education and weight status. *International Journal of Behavioral Nutrition and Physical Activity, 6*(1), 17.

Thomas, S., Reading, J., & Shephard, R. J. (1992). Revision of the physical activity readiness questionnaire (PAR-Q). *Canadian Journal of Sports Science, 17*(4)*,* 338-345.

Troiano, R. P., Berrigan, D., Dodd, K. W., Mâsse, L. C., Tilert, T., & Mcdowell, M. (2007). Physical Activity in the United States Measured by Accelerometer. *Medicine & Science in Sports & Exercise*, *40*(1), 181-188.

Trost, S. G., Pate, R. R., Freedson, P. S., Sallis, J. F., & Taylor, W. C. (2000). Using objective physical activity measures with youth: how many days of monitoring are needed? *Medicine and Science in Sports and Exercise*, *32*(2), 426-431.

Tudor-Locke, C., Johnson, W. D., & Katzmarzyk, P. T. (2009). Accelerometer-Determined Steps per Day in US Adults. *Medicine and Science in Sports and Exercise*, *41*(7), 1384.

US Department of Health and Human Services (USDHHS). (2008). *Physical Activity Guidelines Advisory Committee Report*. Washington, DC: USDHHS, p. vi-14.

Warburton, D. E., Nicol, C. W., & Bredin, S. S. (2006). Health benefits of physical activity: the evidence. *CMAJ*, *174*(6), 801-809.

Ward, D. S., Evenson, K. R., Vaughn, A., Rodgers, A. B., & Troiano, R. P. (2005). Accelerometer use in physical activity: best practices and research recommendations. *Medicine and Science in Sports and Exercise, 37*(11), S582.

Welk, G. J., Blair, S. N., Wood, K., Jones, S., & Thompson, R. W. (2000). A comparative evaluation of three accelerometry-based physical activity monitors. *Medicine and Science in Sports and Exercise*, *32*, S489-97.

Young, D. R., Jerome, G. J., Chen, C., Laferriere, D., Vollmer, W. M. (2009). Patterns of physical activity among overweight and obese adults. *Preventing Chronic Disease, 6*, 1-9.

Chapter 7

VALIDITY OF THE MOVEMENT AND ACTIVITY IN PHYSICAL SPACE (MAPS) SCORE IN HEALTHY ADULTS

Abstract

The ability to accurately assess an individual's physical activity is important as nearly 50% of US adults are insufficiently active at health enhancing levels. Understanding where individuals engage in physical activities may provide information to offer effective ways to promote physical activity behavior change. The purpose of this study was to evaluate validity and reliability evidence of a new method to quantify one's person-environment interaction called the Movement and Activity in Physical Space (MAPS) score. A cross-sectional study was designed to evaluate the distribution of MAPS scores in 75 healthy adults (n = 38 females, n = 37 males; age = 40.4 ± 13.8 yrs; percent fat = 28.9 ± 9.5%). Participants completed three days of monitoring while wearing an ActiGraph accelerometer and LandAirSea GPS receiver on a waist belt during all waking hours. Mean daily belt wear time was 846 ± 149 minutes. To assess validity, Pearson Product Moment Correlation was used to compare MAPS with assessments of physical activity, self efficacy, self-rated health, and neighborhood environment. Reliability was evaluated using Intraclass Correlation Coefficient (ICC). All analyses were performed in SAS 9.2 with alpha level 0.05. Overall MAPS scores trended higher in

males (80.7 ± 60.9) than females (71.9 ± 45.6) with a combined mean of 76.5 ± 53.9 and range 266.0 points. Participants took 6.7 ± 2.8 trips/day (Males = 6.9 ± 2.9; Females = 6.4± 2.8). Minutes of moderate (r = 0.54) and vigorous (r = 0.29) intensity activity from the accelerometer and self rated general health by the SF36 (r = 0.27) were positively associated with MAPS scores. Other measures were not significantly related. The three day reliability of the MAPS score was 0.81 (95% CI = 0.70-0.88). This study demonstrated limited validity and good reliability evidence of MAPS scores in a healthy adult population, which provides the groundwork for future research to assess person-environment interaction in physical activity related studies.

Introduction

Physical activity promotion is a public health priority and the benefits of physical activity and a physically active lifestyle are well documented (Haskell, Lee, et al., 2007; Kohl, 2001; Lee, 2010; Pate, Pratt, Blair, et al., 1995; Tudor-Locke, Brashear, Johnson, & Katzmarzyk, 2010). Despite this, fewer than half of adults engage in physical activities at levels recommended to lower risks for premature morbidity and mortality (Brock et al., 2009; Carlson, Densmore, Fulton, Yore, & Kohl, 2009; Troiano et al., 2007). Survey data from the three different national surveillance systems (i.e., National Health Interview Survey, National Health and Nutrition Examination Survey, and Behavioral Risk Factor Surveillance System NHANES) identify 30% - 48% of adults engage in health enhancing physical activity (Carlson et al., 2009). Findings from the 2003-2004 NHANES accelerometer study identify fewer than 5% of adults engage in physical activity at recommended levels (Troiano et al., 2007). Despite various methods used to assess physical activity, fewer than half of US adults are sufficiently active for optimal health.

Various strategies have been used to help adults increase their physical activity levels. These include individual theory-based interventions and macro-level interventions that address policy and environmental attributes of a community (Fjeldsoe, Marshall, & Miller, 2009; Kinmonth et al., 2008; Wilcox et al., 2007). The ideal intervention may be one that combines strategies to help understand how individuals interact with their environment and provide targeted interventions to change behaviors. This person-environment interaction method would integrate physical activity assessment methods to include pedometers, accelerometer, physical activity logs and/or records, Global Positioning Systems, and audits of environmental and policy features of a community. Pedometers and accelerometers are common objective instruments for assessing physical activity intensity and duration; however these tools do not provide contextual information about a person such as location, distance travelled and speed. Recently, Maddison and Mhurchu (2009) described potential benefits of using the Global Positioning System (GPS) for the ability to include contextual information in physical activity assessment which may enhance understandings of physical activity and sedentary behaviors.

Researchers have found that GPS is a useful tool to compliment traditional physical activity measures in a variety of settings (Maddison & Mhurchu, 2009; Rodríguez, Brown, & Troped, 2005). GPS has been utilized in healthy populations to better understand the impact of the built environment on physical activity (Troped, Wilson, Matthews, Cromley, & Melly, 2010), the use of parks (Quigg, Gray, Reeder, Holt, & Waters, 2010), transport related physical activity (Oliver, Badland, Mavoa, Duncan, & Duncan, 2010), adolescents active travel to schools and during outdoor play (Cooper, Page, Wheeler, Griew, et al., 2010; Cooper, Page, Wheeler, Hillsdon, et al., 2010), in populations to assess functional recovery from surgery (Herrmann et al., in press), and monitoring function in individuals with multiple sclerosis (MS) (Snook, Scott, Ragan, Morand, & Sackau, 2010).

The Movement and Activity in Physical Space (MAPS) score was recently developed to incorporate data from GPS and accelerometers to provide an outcome measure to evaluate free-living function (Herrmann & Ragan, 2008). The MAPS score is based on the definition of function described by the World Health Organization’s International Classification of Functioning, Disability and Health (ICF) as the dynamic interaction of a person’s physical activity within his or her environment (Schneidert, Hurst, Miller, & Ustün, 2003; Ustün, Chatterji, Bickenbach, Kostanjsek, & Schneider, 2003; World Health Organization, 2001). Two devices are used to create the MAPS score, an accelerometer that provides a measure of physical activity and a GPS receiver that assesses mobility within the environment. Combining these data provides a standardized index of free-living function referred to as a MAPS score.

To date, MAPS scores have been used to assess recovery following knee surgery and to monitor function in persons with multiple sclerosis (MS) (Herrmann et al., in press; Snook, Scott, Ragan, Morand, & Sackau, 2010). Herrmann et al. (in press) used MAPS scores to evaluate free-living function in a small sample of patients following knee surgery over a two month period. MAPS scores were shown to accurately identify the decreased level of function and were more sensitive in identifying change over time compared to a subjective functional measure. MAPS scores were considered reliable over three days with Intraclass Correlation Coefficient (ICC) ranging from 0.68 – 0.81. Since there were no objective outcome measures to assess real world functioning in patients with MS, Snook et al. (2010) used MAPS scores to monitor patients for five days. Significant differences were observed in MAPS scores and essential versus non-essential trips by severity of MS and MAPS was more sensitive than accelerometer data alone.

These results provide a foundation of knowledge for MAPS, indicating it may be a viable method for understanding physical activity and sedentary behaviors and free-living function. This measure may have additional utility in other populations and to evaluate other research questions. However, to date there is little data on MAPS scores in a healthy population and the association with other activity related measures is lacking. The purpose of this study was to evaluate the convergent validity of MAPS with measures of physical activity and correlates of physical activity behaviors and to assess the reliability of MAPS scores in 75 healthy adults.

Methods

Participant Selection and Study Design

This study was a cross-sectional design to evaluate MAPS scores in 75 healthy adults (*n* = 37 males, *n* = 38 females). Participants were recruited from a large metropolitan area in the Southwestern United States and screened using the physical activity readiness questionnaire (PAR-Q). Seventy-Five volunteer participants (n=38 females, n=37 males; age = 40.4 ± 13.8 yrs; percent fat = 28.9 ± 9.5%) were enrolled in this study between November 2009 and May 2010. Each participant signed an informed consent form approved by the university's intuitional review board during the first visit and measures of height, weight, and body composition were obtained. Participants were compensated with $50 for successful completion of the study.

Descriptive Measures

Age in years, sex, height in centimeters (cm), and weight in kilograms (kg) were assessed during a visit with study staff. Height was measured three times with participants standing in their bare feet using a Seca portable stadiometer (RoadRod, Hamburg Germany). The average of the two closest readings was used. Weight was measured once using a Tanita scale (Model TBF-300A; Tanita Corporation: Arlington Heights, IL). Body mass index (BMI) was computed as weight in kg divided by height in meters^2.

Objective Measures

The ActiGraph model GT3X accelerometer (ActiGraph, LLC, Pensacola, Florida, USA) was used to record physical activity. The GT3X is a solid state triaxial sensor (micro-electro-mechanical systems) accelerometer (3.8cm x 3.7cm x 1.8cm; 27 grams) that assesses motion in three axes to provide information about physical activity intensity and duration. Worn at the waist level, the GT3X records accelerations ranging in magnitude from approximately 0.05 to 2.5G with a digital filter frequency response from 0.25 to 2.50 Hz. Output data are digitized at a rate of 30 times per second with intensity data recorded in one minute epochs (sampling interval) for this study. Data were downloaded using ActiGraph software and stored in a computer database.

LandAirSea Tracking Key Pro GPS receiver was used to record latitude, longitude, and altitude to provide data on location, speed, trip time, and trip distance. The GPS device (7.6cm x 5.0cm x 3.6cm; 56.7 grams) records data every second and is capable of logging up to 100 hours of movement data with internal flash memory and is powered by two AA batteries. The receiver has 16 parallel channels that continuously track and use up to 16 satellites to compute and update position information for a horizontal accuracy of <3 meters. The receiver acquires GPS position in 39 seconds from a cold start, 30 seconds from a warm start, and < 3 seconds from a hot start (manufacturer information). A motion detector puts the device into a sleep mode following two minutes of no movement to conserve battery life. The data are downloaded via USB to a computer and then, using the LandAirSea software, can be viewed as a text log, displayed over an animated digital street map, or displayed over satellite imagery using Google Earth. A travel logbook was used to confirm the

GPS/GIS assessed locations and allowed participants to record their main activity in each location (e.g., eating, sitting, shopping, etc).

Self-report Measures

The Neighborhood Health Questionnaire (NHQ) is comprised 28 items in seven dimensions (i.e., aesthetic quality, walking environment, availability of healthy foods, safety, violence, social cohesions, and activities with neighbors) (Mujahid, Roux, Morenoff, & Raghunathan, 2007). Neighborhood was defined for the participants as the area approximately 1 mile (1.6km) around their home. Response options are in the form of a Likert scale with a mean score calculated for each domain.

The Global Physical Activity Questionnaire (GPAQ) is a 16 item physical activity questionnaire currently used by the World Health Organization in approximately 50 countries (Armstrong & Bull, 2006). The GPAQ assesses moderate and vigorous physical activity for work, travel to and from places, recreational activities, and time spent in inactivity. The GPAQ takes about five minutes to administer and can be scored for average activity in $\mathrm{min{\cdot}d^{-1}}$ or MET-$\mathrm{min{\cdot}d^{-1}}$.

Self-efficacy and social support for exercise was assessed separately on self-administered five item questionnaires (Eyler et al., 1999; Marcus, Selby, Niaura, & Rossi, 1992). Self-efficacy items were scored using a Likert scale with total possible scores ranging from 5 to 35. Social support items were scored on a Likert scale with total possible scores ranging from 5 to 20. Higher scores indicate greater confidence and greater social support.

A travel diary was used by participants to record the locations they visited throughout the three days of monitoring. Participants recorded the times they would arrive and depart locations and the mode of transportation they used for travel (e.g., walk, bike, bus, car, etc.).

Medical Outcomes Study Short Form 36 (SF-36) includes 36 items over eight domains (i.e., physical functioning, social functioning, role limitations due to physical or emotional problems activities, bodily pain, general mental health, vitality, and general health) (Ware & Sherbourne, 1992). Scores are calculated for each domain with a higher score being desirable. There is a substantial amount of literature to support the use of the SF-36 with over 10,000 citations in published literature.

Procedures

Participants made two visits to the study office to complete study materials. During the first visit, each participant read and signed an informed consent form, completed study questionnaires and had their height and weight measured. The participants were then instructed to wear two devices, an accelerometer and a GPS receiver, on a waist belt during all waking hours and complete a detailed travel diary over for the following three days (Rodríguez, Brown, & Troped, 2005). During the second visit, four to seven days later, participants returned the devices and completed the remaining study questionnaires.

Data Management

For this study, accelerometer non-wear periods were identified as 60 consecutive minutes with no movement data (zero counts) allowing up to 2 consecutive minutes of 1-100 cts$\cdot$min^{-1}. Non-wear periods were ended with >100 cts$\cdot$min^{-1} or with 3 consecutive 1–100 cts$\cdot$min^{-1} (Matthews, Chen, et al., 2008). Accelerometer measured time spent in each intensity level was determined using a SAS Statistical Program. Cut points were used to identify time spent in inactivity (<100 cts/min), light intensity (100-1951 cts/min), moderate (1952–5724 cts/min) and vigorous intensity PA (5725+ cts/min) (Freedson, Melanson, & Sirard, 1998; Matthews et al., 2008; Swartz et al., 2000).

Physical activity and geospatial data were combined to create MAPS scores. The MAPS formula was created to incorporate measures of activity, time, and location data from GPS, geographic information systems (GIS) and from accelerometers to produce a single composite score (Herrmann & Ragan, 2008). A higher MAPS score indicates a higher level of physical activity and environmental interaction based on a combination of more activity and greater environment interaction.

The formula for MAPS is:

$$MAPS = \sum_{L=1}^{n} \left(\frac{Activity}{Minutes} \right) \quad (1),$$

Where L represents locations other than home (environment interaction). The numerator, activity, is a measure of physical activity volume using the step count feature of an accelerometer or a measure of activity intensity determined from an accelerometers activity counts. The accelerometer provides minute-by-minute data on activity and step count. We only used the accelerometer step count data in this study. The denominator, minutes, equals the number of minutes spent at each location for greater than 10 minutes. Ten minutes was chosen as the minimum time spent at a location to be an eligible location. This was identified following pilot testing to reduce confusion and misclassifying of locations due to long stoplights, restaurant drive thru, traffic congestion, and other incidental/unintentional stops. Locations where participants spent less than 10 minutes were not included in the MAPS score calculation. Travel time was included in the MAPS score if a participant used a form of active travel such as walking, biking or bicycling. Table 1 provides a brief example of MAPS data collected from one of the participants (identified as participant ‘X’) during part of a day. Figures 1 and 2 depict the same information from table 1 in a visual map format.

Table 1

Partial Day MAPS Data from Participant X

Location visited	Arrival Time	Departure Time	Duration (min)	Steps	MAPS score
3) Friends House	12:26	14:00	94	511	5.4
4) Shopping	14:11	14:48	37	706	19.1
5) Grocery Store	14:54	15:56	62	1455	23.5

Note. Time is on a 24 hour clock. Daily MAPS scores are summed from each location.

.

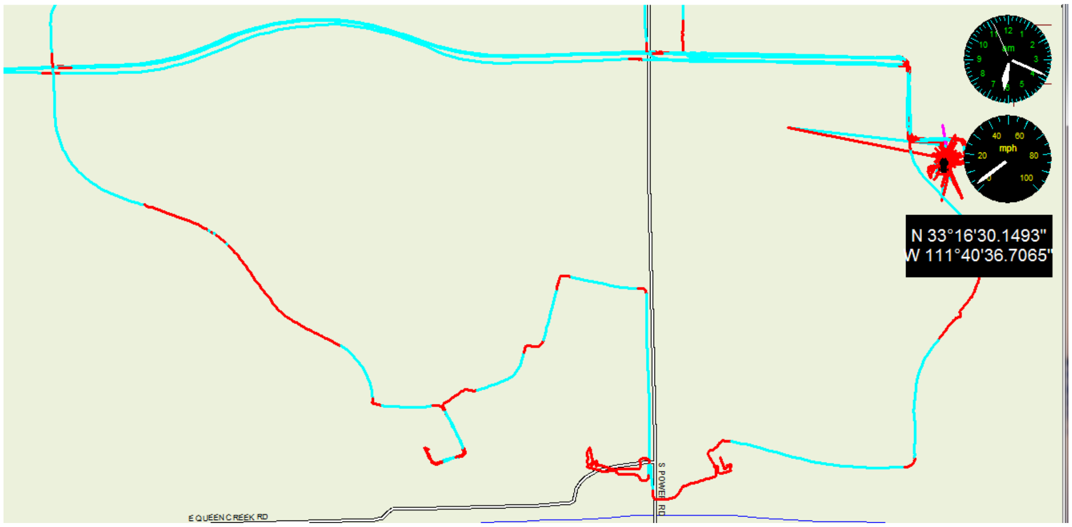


Figure 1. Partial GPS data from Participant X displayed on a digital street map. The blue and red lines represent the travel path with red indicating slow speeds.

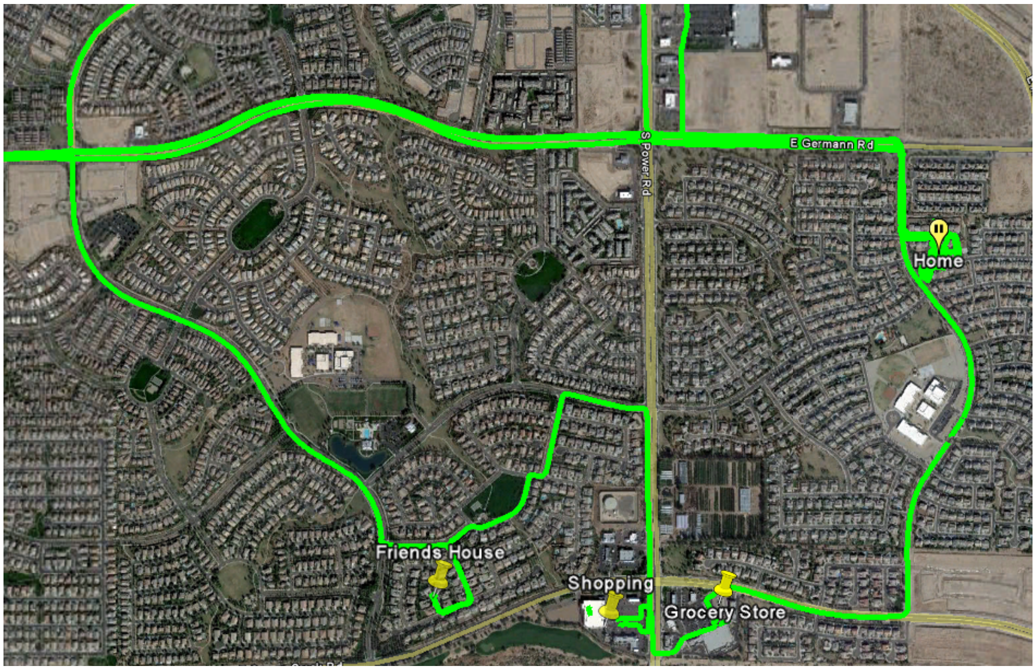


Figure 2. Partial GPS data from Participant X displayed on a satellite image with the travel pathway (green line) and locations (push pin icons)

identified for home, a friend’s house, a shopping destination, and a grocery store.

Data Analysis

Data were analyzed to identify the descriptive profile of the participants, reliability of the MAPS scores over 3 days, and the validity of the MAPS scores against a set of measures that are related to physical activity. Descriptive statistics including mean and standard deviation were computed for demographic data of age, sex, height, weight and BMI, accelerometer data for minutes spent in inactivity, light-, moderate-, and vigorous- intensity activity, GPS measured trips/day, and MAPS scores. Tertiles for MAPS scores were computed to show the variation in participants classified as low, middle, and high for their activity-environment interaction. Analysis of Variance was used to assess differences in MAPS scores by sex, age, BMI, physical activity, NHQ, GPAQ, and Social Support and Self Efficacy for exercise scores. The reliability of MAPS scores was evaluated using an Intraclass Correlation Coefficient (ICC). All analyses were performed in SAS 9.2 with alpha level 0.05.

Results

Data were lost from four participants due to accelerometer malfunction/loss of data (1), GPS malfunction/loss of data (2), and non compliance (1). Table 2 displays descriptive data from the 71 participants completing the study showing means and standard deviations for men, women, and the total sample for the combined MAPS scores, GPS trips/day, accelerometer steps/day and time spent in inactivity, light-, moderate-, and vigorous-intensity physical activity per day. MAPS scores ranged from 5.7 to 271.7 with a mean of 75.1± 55.1. Overall MAPS scores were trending higher in males than females although the difference was not statistically significant ($p > 0.05$). The number of trips/day and time spent in different movement intensities were similar between males and females ($p > 0.05$). Average accelerometer/GPS receiver wear time was 846 ± 149 minutes.

Table 2

Descriptive Data of Study Sample by Sex (Mean ± SD)

	Total (n = 70)	Men (n = 36)	Women (n = 34)
MAPS			
MAPS score	76.5 ± 53.9	80.7 ± 60.9	71.9 ± 45.6
GPS Receiver			
Locations·d^{-1}	6.7 ± 2.8	6.9 ± 2.9	6.4 ± 2.8
Accelerometer			
Steps·d^{-1}	7,792 ± 3,476	8,520 ± 3691	7,064 ± 3,128
Inactivity (min·d^{-1})	534.8 ± 154.8	520.3 ± 164.3	549.3 ± 144.1
Light (min·d^{-1})	271.3 ± 91.2	280.5 ± 97.3	262.2 ± 84.1
Moderate (min·d^{-1})	34.4 ± 33.6	37.2 ± 38.5	31.7 ± 27.8
Vigorous (min·d^{-1})	5.4 ± 13.4	6.7 ± 15.5	4.2 ± 10.8
Demographic			
Age (yr)	40.4 ± 13.8	40.2 ± 13.0	40.1 ± 14.9
Height (cm)	170.9 ± 9.7	178.1 ± 7.3	163.9 ± 5.9
Weight (kg)	78.6 ± 17.7	87.2 ± 13.9	70.2 ± 17.1
BMI (kg/m^2)	26.8 ± 5.7	27.4 ±3.9	26.3 ± 7.0

Notes. MAPS = Movement and Activity in Physical Space; GPS = Global Position System; BMI = Body Mass Index.

Table 3 shows the descriptive data and the accelerometer data by tertiles of MAPS scores. MAPS scores were stratified by tertile ranges (lowest third mean = 25.8 MAPS, middle third mean = 66.5 MAPS, and highest third mean = 137.1 MAPS).

Table 3

Selected Variables by MAPS Tertiles (Mean ± SD)

	Low	Middle	High
MAPS score	25.8 ± 12.1	66.5 ± 13.2	137.1 ± 45.0
GPS Receiver			
Locations$\cdot$d^{-1}	5.1 ± 1.7	7.4 ± 1.8	7.7 ± 1.7
Accelerometer			
Steps$\cdot$d^{-1}	6,112 ± 2,005	7,482 ± 2,300	10,828 ± 3,506
Inactivity (min$\cdot$d^{-1})	565.9. ± 134.9	613.2 ± 139.3	556.8 ± 100.3
Light (min$\cdot$d^{-1})	308.7 ± 94.7	304.4 ± 79.2	272.6 ± 71.0
Moderate (min$\cdot$d^{-1})	19.4 ± 11.7	29.9 ± 14.6	62.5 ± 45.0
Vigorous (min$\cdot$d^{-1})	1.8 ± 6.9	4.3 ± 8.0	9.7 ± 18.7
Demographic			
Age	42.9 ± 13.52	36.1 ± 13.5	42.6 ± 13.6
BMI	27.1 ± 6.1	27.1 ± 4.7	27.1 ± 6.6
% males	47.8%	52.1%	56.5%

Notes. MAPS = Movement and Activity in Physical Space; GPS = Global Position System; BMI = Body Mass Index.

Validity

There were significant correlations between MAPS scores and time spent in moderate ($r = 0.53$, $p < 0.01$) and vigorous ($r = 0.29$, $p = 0.02$) intensity specific categories measured by the accelerometer. There was a borderline inverse association with light intensity activity ($r = 0.25$, $p = 0.05$) and no association with inactivity time ($r = -.03$, $p = 0.85$).

Also, there were no associations between MAPS scores and GPAQ reported minutes of moderate or vigorous intensity activity or total MET minutes per week accumulated in physical activity (all $p > 0.05$). MAPS scores were not significantly associated with domains of the NHQ ($p > 0.05$). Social support and self-efficacy for exercise were not associated ($p > 0.05$) with MAPS scores ($r = 0.16$ and $r = 0.11$ respectively). MAPS scores had a significant correlation with general health measured by the SF36 ($r = 0.27$, $p = 0.03$) but was not associated with other domains (all $p > 0.05$).

Table 4

Correlations Between MAPS Scores and Study Variables

Study Variable	MAPS Scores (*r*)
Age	-.01
BMI	-.04
ActiGraph Accelerometer ($min \cdot d^{-1}$)	
Inactivity	-.03
Light	-.26
Moderate	.54*
Vigorous	.29*
Global Physical Activity Questionnaire ($MET\ min \cdot wk^{-1}$)	
Leisure	-.08
Occupation	-.11
Transportation	.11
Total	-.05
Sitting ($min \cdot d^{-1}$)	.13
Neighborhood Health Questionnaire	
Walk Environment	-.14
Aesthetics	.12
Healthy Foods	.11
Safety	.10
Social Cohesion	.08
Activities with neighbors	.04
Violence	-.11
Medical Outcomes Study Short Form-36	
General Health	.27*
Limitations due to Physical Problems	.20
Social functioning	.19
Pain	.18
Physical Functioning	.10
Emotional Wellbeing	.09
Limitations due to Emotional Problems	.08
Energy/Fatigue	-.01
Social Support for Exercise	-.17
Self Efficacy for Exercise	.11

Note. * $p < 0.05$

Reliability

Three days of monitoring using MAPS scores indicated acceptable reliability indicated acceptable reliability (acceptable > 0.7). The three day reliability of the MAPS score was 0.81 (95% CI = 0.70-0.88).

Discussion

This study used a novel approach to assess person-environment interaction using MAPS scores to combine data from accelerometers and GPS receivers. The MAPS formula was recently developed to provide an index to compare the free-living function as an outcome measure (Herrmann et al., in press). This measure may have additional utility to assess a variety of personal behavior patterns and the interaction with the environment.

Results from the current study indicate that MAPS scores are stable in a healthy population with an ICC = 0.81. This is greater than 0.70 which is considered acceptable for physical performance measures (Nunnally & Bernstein, 1994; Rowe, Mahar, Raedeke, & Lore, 2004) and near 0.80 that is often used in physical activity assessment with accelerometers and pedometers (Matthews, Ainsworth, Thompson, & Bassett, 2002; Tudor-Locke et al., 2005). Our reliability estimate is similar to a previous report using three days of MAPS scores in a small sample that showed ICC values ranging from 0.75 to 0.78 in nine post-surgical knee patients and 0.68 to 0.81 in nine healthy control participants (Herrmann et al., in press).

The value of using a MAPS score is that it integrates both movement data and location/travel data. We used step counts obtained from an ActiGraph accelerometer to represent the amount of movement that occurred at each location our participants traveled. Attainment of a higher step count at each location reflects a higher intensity of movement. Indeed, we observed significant correlations between moderate and vigorous intensity scores on the order of r = .29 to .54 and a nearly significant score for light-intensity activity (r = .05) obtained from the accelerometer and the MAPS score for our population. This suggests the MAPS score is a good measure of activity occurring at various locations traveled. Interestingly, there were no associations between the ActiGraph minutes for inactivity and the MAPS score. The lack of association between inactivity time and MAPS scores likely reflects the large amounts of inactivity that occurs at home. Activity and sedentary behaviors performed at home are not included in MAPS scores because MAPS was developed to assess activity and environment interactions outside of the home.

To assess the validity evidence of an instrument, several types of validity evidence were sought, including convergent and discriminant validity. Convergent validity is illustrated when measures of the same construct are moderately-to-highly correlated, while discriminant validity is characterized by measures of different constructs that do not correlate highly (Rowe & Maher, 2006). We identified demographic characteristics of age and BMI and several instruments that have shown to be statistically associated with physical activity participation to include scores on scales for self-efficacy for exercise, social support for exercise, physical health, general well being, and neighborhood attributes believed to be supportive for physical activity. Results from correlation analysis showed very weak associations with age, BMI, the GPAQ, social support and self efficacy for exercise, and the NHQ (all r < .17). Somewhat surprising, the only score significantly related to the MAPS score was general health, obtained from the SF-36 scale. Perhaps people who perceive their health to be better are more physically activity at locations they visit and visit more locations outside of their home. It is possible that the scales selected, such as self-efficacy and social support for exercise, aspects of a healthy neighborhood, and physical health were measuring different constructs than the MAPS score. For example, some study participants traveled places to engage in exercise, while others did not exercise at all. The locations traveled varied widely from driving to work or school, going to the store, visit family and friends, and to engage in exercise. Thus, the low correlations observed in this study should not discount the validity of the MAPS scores, but rather a need to identify potential correlates of travel and location based behaviors as more appropriate convergent measures for future validity studies.

We thought we would find positive associations between the MAPS and the Global Physical Activity Questionnaire scores since MAPS includes a component of physical activity. The assumption was that people with higher MAPS scores would be more physically activity and report higher levels of domain-specific physical activity and total physical activity. This was evident by accelerometer data when the scores were displayed by tertiles. The GPAQ, a self-report measure, has been correlated with objective measures (pedometers and accelerometers) designed to assess the same construct (i.e., physical activity) and has demonstrated poor to fair associations ranging from r = .06 to .35 (Bull, Maslin, & Armstrong, 2009). However, these correlations are low suggesting weak association at best. As with many self-report questionnaires, the GPAQ may suffer from a certain amount of recall bias resulting in inaccurate estimates of physical activity.

We also thought we would find a stronger association between MAPS scores and the NHQ since studies show positive associations between the built environment and physical activity. Numerous studies have investigated the built environment and oftentimes rely on self-report audits and survey measures, such as the NHQ (Mujahid, Roux, Morenoff, & Raghunathan, 2007). It is possible that the NHQ was not correlated with MAPS scores, because the NHQ relies on the evaluation of the neighborhood within 1 mile of home and MAPS reflects physical activity and interaction with the environment wherever people travel. Thus, the sources of data between the NHQ and MAPS scores, while occasionally overlapping, are measuring different areas where physical activity is performed. The strength of MAPS scores is that it includes GPS data which permits an objective identification of the amount of time people spend travelling to and spending at a locations both inside and outside of their community or neighborhood environment.

The inclusion of MAPS scores has the potential to strengthen associations about person-environment interactions in studies reported in the literature. Recently, Berke et al. (2007) studied the association between the built environment with physical activity and obesity in older adults. The researchers used extensive geographic coding data (e.g., land use, parks, trails, etc) for 1 and 3 km buffer zones around participants' houses and used a walkability audit for each neighborhood. The measure of physical activity was done with only one questions "During the last year, how many days per week did you walk for exercise for at least 15 minutes at a time?" (Berke, Koepsell, Moudon, Hoskins, & Larson, 2007). Their study found that neighborhood characteristics were associated with frequency of walking. However, they used a self report measure of walking frequency that did not specify walking within ones neighborhood, casting some doubt on their conclusion. While study designs such as the one reported by Berke et al (2007) are common and provide insight into the importance of the environment and engagement of physical activity behaviors, inclusion of direct measures person-environment interaction with a MAPS score would strengthen types of information related to both location and activity within geographic data.

Two other studies highlight potential uses for a MAPS score. The International Physical Activity and Environment Network (IPEN) study was designed to help standardize research protocols investigating physical activity and environmental relationships (Badland et al., 2009). The IPEN encourages researchers to use standard methods for evaluating the built environment using geographic information systems and standard methods for physical activity assessment (i.e., accelerometry and the International Physical Activity Questionnaire [IPAQ]). However, the ability to show that the physical activity being assessed is taking place in the neighborhood or community of interest is limited by the instruments used. In the European Union ALPHA study, Spittaels et al., (2010) used survey methodologies to develop and evaluate a new questionnaire to assess European neighborhood environments. They identified this new instrument as valid after using accelerometers and the IPAQ to assess physical activity levels (Spittaels et al., 2010). The researchers found that time spent in various physical activity intensities (light, moderate, vigorous) was associated with some of the survey measured neighborhood environment domains. However, several associations were difficult to interpret, for instance, why was “neighborhood aesthetics” associated with light intensity physical activity, while “safety from traffic” was associated with vigorous intensity activity? Neither the physical activity survey (IPAQ) nor accelerometer

(ActiGraph 7164) used in their study indicate whether the physical activity actually occurred within the neighborhood of interest. Thus, it is difficult to interpret and make use of their results. Using GPS in combination with an objective measure of physical activity to identify where activity occurs may provide more meaningful information about the validity of an instrument designed to assess the importance of neighborhood environments for physical activity.

The studies described above demonstrate some of the added value of using MAPS. The MAPS score is a novel method for objectively assessing person-environment interaction. MAPS does not just measure physical activity or where someone is, but includes data from both accelerometers and GPS to better understand how these two components go hand in hand. As shown in this current study, MAPS has good reliability over three days and is positively related to the user’s daily activity in moderate and vigorous intensity physical activity and one’s rating of their general health.

To date, MAPS scores have been used to assess free-living function in patients during recovery from knee surgery (Herrmann et al., 2008; Herrmann et al., in press) and in patients with MS (Snook et al., 2010). The hypothesis is that people who are healthier interact more with their environment and engage in more activity while outside of the home. Herrmann et al (in press) monitored nine post-surgical knee patients for three days within one week of surgery and again two months later. Initial

MAPS scores were low (14.2 ± 9.7) and increased significantly two months later (35.6 ± 13.9) corresponding with recovery (Herrmann et al., 2008; Herrmann et al., in press). The initial MAPS scores indicated significant functional limitations (due to pain, using crutches, etc.) affecting physical activity and the ability to activity interact within the environment. The mean MAPS score at two months was slightly higher than that of the lowest tertile group in our study (25.8) and indicated the patients had achieved a higher level of function but still may have been limited in certain aspects. This disability study also included a group of nine healthy matched controls that displayed higher MAPS scores that remained stable from the initial measure (57.3 ± 32.5) through the two month follow-up (62.3 ± 29.6) (Herrmann et al., 2008; Herrmann et al., in press). The MAPS scores from the healthy group were slightly lower than the mean MAPS scores in the current study of 71 healthy adults (76.5 ± 55.1) and very similar to the middle tertile group (66.5) suggesting these values indicated a normal range of MAPS scores in healthy individuals.

Snook et al. (2010) conducted a case control study in two patients with MS. Snook used MAPS scores to assess free-living function over five days. MAPS scores were markedly lower (5.2) in the patient that had MS for 21 years with moderate-to-severe symptoms (sometimes used an assistive walking device and was on medication to manage symptoms). The other patient had MS for 7 years with mild symptoms and displaying no functional limitations, did not use an assistive device, was not on MS medications, and demonstrated a higher MAPS score (40.4) (Snook et al., 2010). The MAPS score for both of these individuals with MS would place them in the lowest tertile group from the current study. Thus, scores in the lowest tertile group from the current study may identify individuals with some form of limitation, either in physical activity behavior or in their interaction with the environment outside of their home (i.e., unable to go certain places).

Limitations

Increasing validity and reliability evidence of instruments is an important element in any field. In the current study with healthy individuals, MAPS scores was evaluated for convergent and discriminant validity evidence using a variety of scores that may reflect different constructs than the MAPS score. On average, the sample was normal-to-overweight, middle-age, generally physically active, and most were employed. Because of this, the sample was relatively homogeneous in their behavior and physical abilities (e.g., similar BMI levels and no physical disabilities) which limits the external validity to adults with physical ability restrictions, person's with low physical activity, or those with dissimilar characteristics (e.g., older adults, obese). It also may have attenuated the correlations observed in the study.

Some variation in MAPS scores may be due to differences in the weather and other conditions that may have affected physical activity behaviors and travel outside of the home. Our study took place from November through May in the Southwest U.S. which is characterized by temperate climate during that time of year. However, variation in the weather, holidays, work schedules, and other obligations may have altered travel and activity behavior patterns that affect the MAPS scores.

Summary

This study demonstrated the use of objective measures of physical activity and location (accelerometers and GPS receivers) to provide a broad evaluation of a person's activity and interaction within their environment using MAPS scores. We found good reliability over three days and evidence of validity with self-reported health status and moderate- and vigorous-intensity physical activity. This type of measure has potential uses in a wide variety of areas to enhance understanding of interactions between physical activity behaviors and the effect of the built environment.

References

Armstrong, T., & Bull, F. (2006). Development of the World Health Organization Global Physical Activity Questionnaire (GPAQ). *Journal of Public Health*, *14*(2), 66–70.

Badland, H., Schofield, G., Witten, K., Schluter, P., Mavoa, S., Kearns, R., Hinckson, E., et al. (2009). Understanding the Relationship between Activity and Neighbourhoods (URBAN) Study: research design and methodology. *BMC Public Health*, *9*(1), 224.

Berke, E. M., Koepsell, T. D., Moudon, A. V., Hoskins, R. E., & Larson, E. B. (2007). Association of the Built Environment With Physical Activity and Obesity in Older Persons. *Am J Public Health*, *97*(3), 486-492.

Brock, D. W., Thomas, O., Cowan, C. D., Allison, D. B., Gaesser, G. A., & Hunter, G. R. (2009). Association Between Insufficiently Physically Active and the Prevalence of Obesity in the United States. *Journal of physical activity & health*, *6*(1), 1-5.

Carlson, S. A., Densmore, D., Fulton, J. E., Yore, M. M., & Kohl, Harold W, 3rd. (2009). Differences in physical activity prevalence and trends from 3 U.S. surveillance systems: NHIS, NHANES, and BRFSS. *Journal of Physical Activity & Health*, *6 Suppl 1*, S18-27.

Cooper, A. R., Page, A. S., Wheeler, B. W., Griew, P., Davis, L., Hillsdon, M., & Jago, R. (2010). Mapping the walk to school using accelerometry combined with a global positioning system. *American Journal of Preventive Medicine*, *38*(2), 178-183.

Cooper, A. R., Page, A. S., Wheeler, B. W., Hillsdon, M., Griew, P., & Jago, R. (2010). Patterns of GPS measured time outdoors after school and objective physical activity in English children: the PEACH project. *The International Journal of Behavioral Nutrition and Physical Activity*, *7*, 31.

Curran, S. L., Andrykowski, M. A., & Studts, J. L. (1995). Short Form of the Profile of Mood States (POMS-SF): Psychometric information. *Psychological Assessment*, *7*(1), 80-83.

Eyler, A. A., Brownson, R. C., Donatelle, R. J., King, A. C., Brown, D., & Sallis, J. F. (1999). Physical activity social support and middle-and older-aged minority women: results from a US survey. *Social Science & Medicine*, *49*(6), 781–789.

Fjeldsoe, B. S., Marshall, A. L., & Miller, Y. D. (2009). Behavior Change Interventions Delivered by Mobile Telephone Short-Message Service. *American Journal of Preventive Medicine*, *36*(2), 165-173.

Freedson, P S, Melanson, E., & Sirard, J. (1998). Calibration of the Computer Science and Applications, Inc. accelerometer. *Medicine and Science in Sports and Exercise*, *30*(5), 777-781.

Haskell, W. L., Lee, I. M., Pate, R. R., Powell, K. E., Blair, S. N., Franklin, B. A., Macera, C. 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. *Circulation*, *116*(9), 1081-1093.

Herrmann, S. D., & Ragan, B. G. (2008). Outcome Assessment Part 2: Measurement of Person-Environment Interaction. *Athletic Therapy Today*, *13*(2), 29-33.

Herrmann, S. D., Snook, E. M., Kang, M., Scott, C. B., Mack, M. G., Dompier, T. P., & Ragan, B. G. (in press). Development and

Validation of a Movement and Activity in Physical Space (MAPS) Score as a Functional Outcome Measure. *Archives of Physical Medicine and Rehabilitation*.

Kinmonth, A. L., Wareham, N. J., Hardeman, W., Sutton, S., Prevost, A. T., Fanshawe, T., Williams, K. M., et al. (2008). Efficacy of a theory-based behavioural intervention to increase physical activity in an at-risk group in primary care (ProActive UK): a randomised trial. *The Lancet*, *371*(9606), 41-48.

Kohl, H. W. (2001). Physical activity and cardiovascular disease: evidence for a dose response. *Medicine and Science in Sports and Exercise*, *33*(6 Suppl), S472-483; discussion S493-494.

Lee, I. M. (2010). Physical activity and cardiac protection. *Current Sports Medicine Reports*, *9*(4), 214-219.

Maddison, R., & Ni Mhurchu, C. (2009). Global positioning system: a new opportunity in physical activity measurement. *The International Journal of Behavioral Nutrition and Physical Activity*, *6*, 73.

Marcus, B. H., Selby, V. C., Niaura, R. S., & Rossi, J. S. (1992). Self-efficacy and the stages of exercise behavior change. *Research Quarterly for Exercise and Sport*, *63*(1), 60–66.

Matthews, C. E., Ainsworth, B. E., Thompson, R. W., & Bassett, D. R. (2002). Sources of variance in daily physical activity levels as measured by an accelerometer. *Medicine and Science in Sports and Exercise*, *34*(8), 1376-1381.

Matthews, C. E., Chen, K. Y., Freedson, P. S., Buchowski, M. S., Beech, B. M., Pate, R. R., & Troiano, R. P. (2008). Amount of time spent in sedentary behaviors in the United States, 2003-2004. *American Journal of Epidemiology*, *167*(7), 875-881.

Mujahid, M., Roux, A., Morenoff, J., & Raghunathan, T. (2007). Assessing the measurement properties of neighborhood scales: From psychometrics to ecometrics. *American Journal of Epidemiology*, *165*(8), 858-867.

Nunnally, J., & Bernstein, I. (1994). *Psychometric Theory* (3rd ed.). McGraw-Hill Humanities/Social Sciences/Languages.

Oliver, M., Badland, H., Mavoa, S., Duncan, M. J., & Duncan, S. (2010). Combining GPS, GIS, and accelerometry: methodological issues in the assessment of location and intensity of travel behaviors. *Journal of Physical Activity & Health*, *7*(1), 102-108.

Pate, R. R., Pratt, M., Blair, S. N., Haskell, W. L., Macera, C. A., Bouchard, C., Buchner, D., et al. (1995). Physical activity and public health. A recommendation from the Centers for Disease Control and Prevention and the American College of Sports Medicine. *JAMA: The Journal of the American Medical Association*, *273*(5), 402-407.

Quigg, R., Gray, A., Reeder, A. I., Holt, A., & Waters, D. L. (2010). Using accelerometers and GPS units to identify the proportion of daily physical activity located in parks with playgrounds in New Zealand children. *Preventive Medicine*, *50*(5-6), 235-240.

Rodríguez, D. A., Brown, A. L., & Troped, P. J. (2005). Portable global positioning units to complement accelerometry-based physical activity monitors. *Medicine and Science in Sports and Exercise*, *37*(11 Suppl), S572-581.

Rowe, D. A., Mahar, M. T., Raedeke, T. D., & Lore, J. (2004). Measuring physical activity in children with pedometers: reliability, reactivity, and replacement of missing data. *Pediatric Exercise Science*, *16*(4), 343-354.

Rowe, D. A., & Maher, M. T. (2006). Validity. In T. M. Wood & W. Zhu (Eds.), *Measurement Theory and Practice in Kinesiology*. Champaign, IL: Human Kinetics.

Schneidert, M., Hurst, R., Miller, J., & Ustün, B. (2003). The role of environment in the International Classification of Functioning, Disability and Health (ICF). *Disability and Rehabilitation*, *25*(11-12), 588-595.

Shacham, S. (1983). A shortened version of the Profile of Mood States. *Journal of Personality Assessment*, *47*(3), 305-306.

Snook, E. M., Scott, C. B., Ragan, B. G., Morand, A. M., & Sackau, T. (2010). Combining GPS and Accelerometer Data as a New Functional Outcome Measure for Multiple Sclerosis: A Case Comparison Study of the Movement and Activity in Physical Space

(MAPS) Score. Presented at the mHealth Summit Conference, Washington, DC.

Spittaels, H., Verloigne, M., Gidlow, C., Gloanec, J., Titze, S., Foster, C., Oppert, J. M., et al. (2010). Measuring physical activity-related environmental factors: reliability and predictive validity of the European environmental questionnaire ALPHA, *7*, 48.

Swartz, A. M., Strath, S. J., Bassett, D. R., Jr, O'Brien, W. L., King, G. A., & Ainsworth, B. E. (2000). Estimation of energy expenditure using CSA accelerometers at hip and wrist sites. *Medicine and Science in Sports and Exercise*, *32*(9 Suppl), S450-456.

Troiano, R. P., Berrigan, D., Dodd, K. W., MâSse, L. C., Tilert, T., & Mcdowell, M. (2007). Physical Activity in the United States Measured by Accelerometer. *Medicine & Science in Sports & Exercise*, *40*(1):181-188.

Troped, P. J., Wilson, J. S., Matthews, C. E., Cromley, E. K., & Melly, S. J. (2010). The built environment and location-based physical activity. *American Journal of Preventive Medicine*, *38*(4), 429-438.

Tudor-Locke, C., Burkett, L., Reis, J. P., Ainsworth, B. E., Macera, C. A., & Wilson, D. K. (2005). How many days of pedometer monitoring predict weekly physical activity in adults? *Preventive Medicine*, *40*(3), 293-298.

Tudor-Locke, C., Brashear, M. M., Johnson, W. D., & Katzmarzyk, P. T. (2010). Accelerometer profiles of physical activity and inactivity in normal weight, overweight, and obese U.S. men and women. *The International Journal of Behavioral Nutrition and Physical Activity*, *7*, 60.

Ustün, T. B., Chatterji, S., Bickenbach, J., Kostanjsek, N., & Schneider, M. (2003). The International Classification of Functioning, Disability and Health: a new tool for understanding disability and health. *Disability and Rehabilitation*, *25*(11-12), 565-571.

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

Wilcox, S., Laken, M., Bopp, M., Gethers, O., Huang, P., McClorin, L., Parrott, A. W., et al. (2007). Increasing Physical Activity Among Church Members: Community-Based Participatory Research. *American Journal of Preventive Medicine*, *32*(2), 131-138.

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

Chapter 8

THE IMPACT OF ACCELEROMETER WEAR TIME ON PHYSICAL ACTIVITY DATA

Abstract

Current research practice employs wide ranging accelerometer wear time criteria to identify a valid data of physical activity assessment. This variation may limit comparability between studies and requires further investigation to understand the effect of wear time on physical activity data. The purpose of this study is to evaluate the effects of varying amounts of daily accelerometer wear time on physical activity in the 2005-2006 National Health and Nutrition Examination Study (NHANES). One thousand days of accelerometer data from 1000 NHANES participants (age = 38.7 ± 14.3 years; BMI = 28.2 ± 6.7 $\mathrm{kg \cdot m^{-2}}$) were randomly selected to use in the semi-simulation approach. A criterion data set was created using 100 random days with 14 $\mathrm{h \cdot d^{-1}}$ of wear time. Additional samples of 100 days were randomly selected with 10-, 11-, 12-, and 13- $\mathrm{h \cdot d^{-1}}$. These data sets were used as day-to-day comparison to create four semi-simulation data sets (10-, 11-, 12-, 13- $\mathrm{h \cdot d^{-1}}$) form the criterion data set. Differences in time spent in inactivity (<100 $\mathrm{cts \cdot min^{-1}}$), light (100–1951 $\mathrm{cts \cdot min^{-1}}$), moderate (1952–5724 $\mathrm{cts \cdot min^{-1}}$), and vigorous (≥5725 $\mathrm{cts \cdot min^{-1}}$) intensity activity was assessed using repeated measures ANOVA and Absolute Percent Error (APE). There were significant differences for moderate intensity activity between the 14 $\mathrm{h \cdot d^{-1}}$ criterion and 10 $\mathrm{h \cdot d^{-1}}$ and

11 h·d^{-1} and for 10 h·d^{-1} to 13 h·d^{-1} for inactivity and light physical activity ($p<0.05$). APE values increased with the shorter accelerometer wear time (13 h·d^{-1} = 6.6% to 14.1%; 12 h·d^{-1} = 10.2% to 15.2%, 11 h·d^{-1} = 18.6% to 35.5%; 10 h·d^{-1} = 26.2% to 40.3%). These data suggest that researchers using a wear time criteria of 12 h·d^{-1} or less may be underestimating time spent in various activity levels.

Introduction

In 2008, the US government released its first physical activity guidelines for Americans, encouraging 150 min/wk of moderate intensity aerobic activity or 75 min/wk of vigorous-intensity physical activity, or the equivalent combination of moderate and vigorous activities. This followed recommendations established in 2007 by the American College of Sports Medicine (ACSM) and the American Heart Association (Haskell, Lee, et al., 2007) and in 1995 by the ACSM and the U.S. Centers for Disease Control and Prevention (Pate et al., 1995). Despite recommendations such as these, the prevalence of physical activity in the United States among adults remains low (Macera et al., 2005; Matthews et al., 2008; Troiano et al., 2007) placing a great importance on research to increase and evaluate physical activity.

Objectively measuring physical activity with accelerometers has become a popular method for physical activity assessment. Standard measures and methodologies for accelerometer use are paramount to assessing physical activity for appropriately classifying individuals and to accurately monitor changes in physical activity. Therefore, understanding these instruments, their measurement properties, and interpreting the output has been the focus of a great deal research. For example, there has been a significant amount of research regarding the number of days a person must wear an accelerometer to accurately assess habitual physical activity levels. Several researchers suggest 3-7 days is sufficient to reflect usual physical activity patterns in adults (Coleman & Epstein, 1998; Gretebeck & Montoye, 1992; Levin, Jacobs, Ainsworth, Richardson, & Leon, 1999; Matthews, Ainsworth, Thompson, & Bassett, 2002; Tudor-Locke, Burkett, Reis, Ainsworth, et al., 2005). Others recommend that 4-9 days are needed to reflect physical activity patterns in children (Janz, Witt, & Mahoney, 1995; Murray, Catellier, Hannan, et al., 2004; Treuth et al., 2003; Trost, Pate, Freedson, Sallis, & Taylor, 2000). There is some variability in the number of days required depending on the population and outcome of interest, such as the days required to assess sedentary behavior or moderate-to-vigorous physical activity. Despite differences in the recommendations, such studies have improved our understanding of habitual physical activity and have guided countless research studies.

Few studies have recommended the optimal number of $h \cdot d^{-1}$ an accelerometer should be worn to identify a valid day (Banda, Hutto, Feeney, Pfeiffer, et al., 2010; Mâsse et al., 2005). There is a wide range of wear time $h \cdot d^{-1}$ criteria used to identify a valid day. Reports in the literature range from as few as two hours (Boarnet, Forsyth, Day, & Oakes, 2011; Forsyth, Oakes, Lee, & Schmitz, 2009) to limiting the upper range to 16 hours (Slootmaker, Schuit, Chinapaw, Seidell, & van Mechelen, 2009). Other have identified valid days by using different criteria for week days (10 $h \cdot d^{-1}$) versus weekend days (8 $h \cdot d^{-1}$) (Spittaels, Verloigne, Gidlow, Gloanec, Titze, et al., 2010), 60% (Mâsse, et al., 2005) to 75% (Matthews et al., 2002) of awake time, or sample specific criteria (Catellier, Hannan, Murray, et al., 2005; Evenson, 2011). In a study of behavior change in children, Baronowski et al (2011) defined a valid day as having at least 800 minutes (~13.3 hours) of accelerometer wear time (Baranowski et al., 2011). Others have required ≥12 $h{\cdot}d^{-1}$ of wear time (Banda et al., 2010; Chinapaw, Slootmaker, Schuit, van Zuidam, & van Mechelen, 2009; Matthews 2002). Catellier et al (2005) proposed the 70/80 rule which requires 70% of the sample to have accelerometer data and 80% of that observed period becomes the valid day threshold (Catellier, Hannan, Murray, et al., 2005; Evenson, 2011). Still others have normalized the data to 12 hours by inputting data (e.g., 10 $h{\cdot}d^{-1}$ of wear time was changed to 12 $h{\cdot}d^{-1}$ increasing minutes in each intensity level proportionally) (Alzahrani, Ada, & Dean, 2011; Young, Jerome, Chen, Laferriere, & Vollmer, 2009). Ten $h{\cdot}d^{-1}$ of accelerometer wear time is regularly used to identify a valid day of accelerometer data (Colley et al., 2011; Motl et al., 2010; Troiano et al., 2007). However, empirical evidence is lacking to support 10 $h{\cdot}d^{-1}$of wear time or that any other wear time criteria is superior to another. Without data to support a consensus for daily wear time ($h{\cdot}d^{-1}$) criteria, the validity of daily wear time and the comparability of studies using different wear time criteria must be questioned.

Recent efforts have been made to better understand the differences and error associated with using different amounts of accelerometer wear time and how this may impact the time reported for movement in different intensities. Using accelerometer data from adults enrolled in a worksite health promotion project to increase walking, a semi-simulation approach was used identify the optimal wear time an accelerometer should be worn to reflect daily activity. Accelerometer data with 14 h·d^{-1} of wear time were compiled into a criterion sample data set and compared to data sets with fewer hours to remove physical activity data, which created new data sets with less wear time (Herrmann et al., in review). The semi-simulated data sets that were created indicated there were significant differences in minutes of sedentary, light, and moderate intensity activity with less than 14 h·d^{-1}of wear time. Furthermore, a considerable amount of error (Absolute Percent Error >10%) in physical activity at all intensity levels was associated with accelerometer wear time of less than 13 h·d^{-1}. Because these findings were observed in a small, homogenous group of employees with a goal of walking 10,000 steps per day, additional research is needed to further understand the variation associated with wearing an accelerometer for varying h·d^{-1} in a population of adults with diverse physical activity behaviors. The purpose of this study is to evaluate the effects of varying amounts of daily accelerometer wear time on physical activity in a sample of adults participating in the 2005-2006 National Health and Nutrition Examination Study (NHANES).

Methods

Study Design and Participant Selection

NHANES 2005-2006 employs a complex, multistage probability sampling method to obtain a representative sample of the U.S. population. The purpose of NHANES is to assess the health and nutritional status of adults and children in the United States for use in understanding the prevalence and risk factors for diseases. NHANES participants undergo extensive evaluations that include interviews and health examinations. In 2003, as part of the evaluation process, all ambulatory participants >6 years were asked to wear a physical activity monitor (accelerometer). The data files for NHANES 2005-2006 was released in June 2008.

The full 2005-2006 accelerometer and demographic data sets were downloaded from the NHANES website (http://www.cdc.gov/nchs/nhanes/nhanes2005-2006/exam05_06.htm). A SAS statistical program was used to identify outlier values due to accelerometer malfunction, to identify now-wear periods, and time spent in each intensity level (inactivity, light, moderate, and vigorous). Data from adults 18 - 65 years, with accelerometer data (approximately N=4,000 participants), were selected for use in this study.

Instrument

The NHANES uses the ActiGraph model 7164 accelerometer (formerly the CSA/MTI AM-7164, manufactured by ActiGraph of Ft. Walton

Beach, Florida, USA) to assess physical activity. The ActiGraph 6164 (5.08 x 3.81 x 1.27cm) uses a single axis accelerometer that measures vertical accelerations. The output frequency is 0.25 to 2.5 Hz, digitized by an analog to digital converter, and stored in one minute epochs (sampling intervals) (Freedson, 1998). The accelerometer is powered by a three volt coin-cell lithium battery.

NHANES participants were provided verbal and written instructions to wear the accelerometer for seven days during all waking hours. The accelerometer was secured to the waist over the right hip by an elastic belt which included a Velcro pouch for the accelerometer. Exclusions for the NHANES accelerometer assessment include waist girths that are too large for the belt (~80cm elastic), individuals in wheelchairs, and those with recent abdominal surgery.

For this study, non-wear periods were identified as 60 consecutive minutes with no movement data (zero counts) allowing up to 2 consecutive minutes of 1-100 cts·min^{-1} (Matthews, Chen, Freedson, et al., 2008). Non-wear periods were ended with >100 cts·min^{-1} or with 3 consecutive 1–100 cts·min^{-1} (Matthews, Chen, Freedson, Buchowski, et al., 2008). ActiGraph monitors were scored to assess time spent in each intensity level using a SAS statistical program. Time spent in inactivity was identified by cut-points < 100 cts·min^{-1} (Ekelund, Yngve, Sjöström, & Westerterp, 2000; Swartz et al., 2000), and Freedson’s cut points were used to determine time spent in light intensity (100-1951 cts·min^{-1}), moderate intensity (1952-

5724 $cts \cdot min^{-1}$), and vigorous intensity (5725+ $cts \cdot min^{-1}$) (Freedson, Melanson, & Sirard, 1998).

Semi-Simulation Approach

A semi-simulation design was used to create new data sets with different amounts of wear time from an original data set. These semi-simulated data sets can then be compared to the criterion data set to understand the amount of error associated with varying amounts of wear time. The basic premise of the semi-simulation approach is that it considers data characteristics (e.g., real accelerometer wear pattern) to remove data instead of a random data removal approach.

The criterion wear time value was set at 14 $h \cdot d^{-1}$ and a random sample of 200 days was selected from NHANES participants that wore the accelerometer for 14 valid hours (i.e., the accelerometer was worn for at least 40 min for each hour) (Evenson & Terry, 2009). Four additional 200-day samples were randomly selected from individuals that wore the accelerometer for 13-, 12-, 11-, and 10 $h \cdot d^{-1}$. These *comparison data sets* were used as reference for their missing data pattern in the semi-simulation approach.

The semi-simulation approach matches the wear time data pattern in a one day-to-one day comparison to remove data from the criterion 14 $h \cdot d^{-1}$ data set (e.g., 14 $h \cdot d^{-1}$ matched with 13 $h \cdot d^{-1}$, 14 $h \cdot d^{-1}$ matched with 12 $h \cdot d^{-1}$, etc.). For example, a criterion day with a 14 hour wear pattern indicates the accelerometer was worn from 7 a.m. until 9 p.m. and a *comparison* day with a 12 hour pattern shows it was worn from 7 a.m. until 7 p.m. Then the data from 7 p.m. to 9 p.m. from the 14 hour day would be removed creating a *semi-simulated* 12 hour day. This one day-to-one day matching and removing data provides four new 200 day *semi-simulated data sets* of 13-, 12-, 11-, and 10 h·d^{-1} with real world missing data patterns. These new *semi-simulated data sets* can be compared with the criterion 14 h·d^{-1} data set to identify differences in min·d^{-1} spent in varying activity intensity levels. Figure 1 displays the data management procedure for the semi-simulation approach. This approach has been used and described in detail elsewhere (see Kang, Rowe, Barreira, Robinson, & Mahar, 2009 and Kang, Zhu, Tudor-Locke, & Ainsworth, 2005).

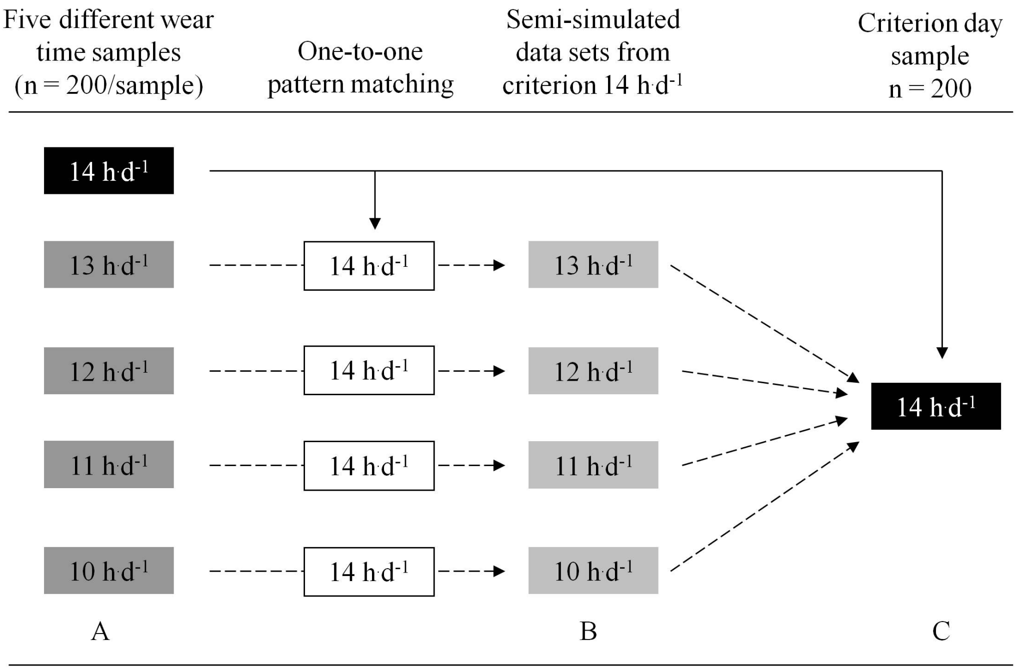


Figure 1. A = h·d^{-1} comparison samples; A-B = Data removed from 14 h·d^{-1} to match pattern of less wear time; and B-C = Compare data of semi-simulated data sets to criterion sample data

Statistical Analysis

Repeated measures ANOVAs with the Least Significant Difference (LSD) post hoc method was performed to assess differences in daily minutes between semi-simulation data sets (13-, 12-, 11-, and 10 h·d^{-1}) and criterion value (14 h·d^{-1}) at each intensity level: inactivity, light, moderate, and vigorous. Absolute Percent Error (APE) was computed between the criterion value and each of the semi-simulation data sets for daily minutes of inactivity, light, moderate, and vigorous activity. APE identifies the percent difference between two values with a lower APE desired (APE = [Observed Value] – [Criterion Value] / Criterion Value *

100). APE has been used in similar analyses to identify the number of days needed to assess habitual physical activity with a pedometer (Kang, Bassett, Barreira, Tudor-Locke, Ainsworth, Reis, et al., 2009). Proportion of time spend in each activity intensity level was calculated for each semi-simulated data set and the criterion value. The alpha level was set at 0.05. SAS version 9.2 (SAS Institute Inc, Cary, North Carolina) were used for the data analysis.

Results

Details of the entire NHANES accelerometer sample can be found elsewhere (see Troiano et al., 2007). The analyzed sample in this study used 1,000 days from 1,000 participants which came from over 6,000 eligible participants that wore an accelerometer for over 30,000 days in total. Demographic data and anthropometric characteristics of the study population are presented in Table 1.

Table 1

Selected Demographic and Anthropometric Characteristics for the Total Study Sample and Criterion Sample

	Total Sample (N = 1,000)			Criterion Sample (n = 200)		
	Total	Men (44.8%)	Women (55.2%)	Total	Men (44.5%)	Women (55.5%)
Age (y)	38.7 ± 14.3	38.1 ± 15.0	39.2 ± 13.8	38.7 ± 14.2	38.1 ± 15.1	29.2 ± 13.8
Height (cm)	167.8 ± 10.5	176.0 ± 8.4	161.1 ± 6.5	167.8 ± 10.5	176.1 ± 8.4	161.2 ± 6.5
Weight (kg)	79.9 ± 22.6	87.4 ± 26.2	73.8 ± 17.0	79.9 ± 22.7	87 .5 ± 26.3	73.8 ± 17.1
BMI (kg/m^2)	28.2 ± 6.7	28.0 ± 7.2	28.4 ± 6.2	28.2 ± 6.7	28.0 ± 7.2	28.4 ± 6.3

Table 2 shows the accelerometer measured minutes of activity by wear time duration (10-, 11-, 12-, 13-, and 14 $h{\cdot}d^{-1}$). Compared to the criterion value of 14 $h{\cdot}d^{-1}$, minutes tended to decrease across activity categories as the wear time decreased. Results from the repeated measures ANOVA showed significant differences in daily minutes between the 14 $h{\cdot}d^{-1}$ criterion value and the semi-simulated data sets of 10-, 11-, 12-, and 13 $h{\cdot}d^{-1}$ for inactivity and light intensity ($p < .05$). There were significant differences for minutes of moderate intensity between the 14 $h{\cdot}d^{-1}$ criterion value and the 10- and 11 $h{\cdot}d^{-1}$ semi-simulated data sets ($p < .05$). No difference was found for vigorous intensity activity ($p > .05$).

Table 2

Mean (SEM) for Minutes of Physical Activity by Intensity and Hours per Day of Accelerometer Wear Time

Activity Level	Referent	Semi-simulated datasets (h·d^{-1})			
	14	13	12	11	10
Inactivity	456.9 (8.1)	421.8 (7.9)*	389.0 (7.4)*	355.6 (6.8)*	322.2 (6.5)*
Light	347.3 (7.5)	326.7 (7.2)*	301.9 (7.0)*	278.2 (6.3)*	253.7 (6.0)*
Moderate	26.3 (2.0)	25.0 (1.9)	23.8 (1.7)	21.0 (1.5)*	19.6 (1.4)*
Vigorous	1.2 (0.4)	1.1 (0.4)	1.1 (0.4)	1.0 (0.4)	0.7 (0.3)

Note. * indicates significance at $p < 0.05$.

Table 3 displays the APE values for accelerometer measured activity by wear time (h·d^{-1}) and activity category (inactivity, light, moderate, and vigorous). APE values ranged from 6.6% for light intensity activity with 13 h·d^{-1} to 40.3% for vigorous intensity activity with 10 h·d^{-1}. APE values increased with the shorter accelerometer wear time (13 h·d^{-1} = 6.6% to 14.1%; 12 h·d^{-1} = 10.2% to 15.2%, 11 h·d^{-1} = 18.6% to 35.5%; 10 h·d^{-1} = 26.2% to 40.3%). There were no h·d^{-1} categories with APE values below 10% for all activity levels.

Table 3

Mean (SEM) Absolute Percentage Error (APE) by Intensity and Hours per Day of Accelerometer Wear Time

Activity Level	Semi-simulated datasets ($h \cdot d^{-1}$)			
	13	12	11	10
Inactivity	7.4% (0.2)	14.9% (0.4)	24.4% (0.5)	29.7% (0.5)
Light	6.6% (0.3)	13.5% (0.4)	20.2% (0.5)	27.5% (0.6)
Moderate	6.9% (1.0)	10.2% (1.5)	18.6% (1.6)	26.2% (1.6)
Vigorous	14.1% (6.0)	15.2% (6.4)	35.5% (8.1)	40.3% (8.3)

Table 4 shows the proportion of time spent in each activity category by $h \cdot d^{-}1$ of wear time. The proportion of time spent in inactivity decreased as wear time decreased with values ranging from 54.2% to 53.2%. The proportion of light intensity activity increased as wear time decreased with values ranging from 42.7% to 43.6%. The proportion of time in moderate and vigorous intensity increased slightly as wear time decreased (moderate 2.9% to 3.1%; vigorous 0.1% to 0.2%).

Table 4

Proportion of Time Spent in Each Activity Level by Hours of Wear Time

Activity Level	Referent	Semi-simulated datasets (h·d^{-1})			
	14	13	12	11	10
Inactivity	54.9%	54.5%	54.3%	54.2%	54.0%
Light	41.8%	42.2%	42.2%	42.4%	42.6%
Moderate	3.2%	3.2%	3.3%	3.2%	3.3%
Vigorous	0.1%	0.1%	0.2%	0.2%	0.1%

Discussion

This study offers insight into the effects of accelerometer wear time on estimates of time spent in varying intensities of physical activity. We applied a semi-simulation approach to a sample of 2005-2006 NHANES participants with accelerometer data to compare the effects of wearing an accelerometer for 10-, 11-, 12-, 13-, and 14 h·d^{-1}. The semi-simulation approach allows for comparison of data from a known 14 h·d^{-1} data by matching the missing accelerometer data pattern with 10-, 11-, 12-, and 13 h·d^{-1} data sets. It also allows researchers to know exactly what data were removed to recreate real life patterns of missing (or non-wear) data to show the impact of reduced accelerometer wear time on time spent in different physical activity intensity levels and inactivity behaviors.

In general, longer wear times provided significantly greater amounts of time in inactivity, light intensity, and moderate intensity physical activity. Furthermore, the amount of error in inactivity and physical activity estimates were greater with less wear time. Few individuals in this sample participated regularly in vigorous intensity activity provided limited information about a highly active sample.

These findings are similar to a previous study we completed using the semi-simulation approach to investigate wear time differences in a sample of adults participating in a worksite health promotion study to increase walking (Herrmann et al., in review). We showed result patterns that were similar to this current study, that with increased accelerometer wear time there were significantly more minutes recorded in inactivity and light and moderate intensity activity. As with the current sample, the duration spent in vigorous intensity activity was insufficient to be affected by accelerometer wear time. However, one prominent difference in the previous study compared to our current analysis was that the previous sample did more moderate intensity activity (45 $min \cdot d^{-1}$) as a result of participating in the walking intervention compared to this sample of NHANES participants (26 $min \cdot d^{-1}$). Despite this large difference, the error patterns remained strikingly similar. Excluding vigorous intensity physical activity, the APE reported for 13 $h \cdot d^{-1}$ ranged from 6.6% to 7.4% for this study and 6.4% to 7.8% from the walking intervention study (Herrmann et al., in review). Both of these studies demonstrate that by reducing the wear time criteria by more than one hour from the 14 $h \cdot d^{-1}$ criterion value, it will result in an unacceptable amount of error (APE > 10%).

These findings have important clinical and research implications for studies using accelerometers. Studies requiring a wear time less than 13 h·d^{-1} may be underestimating the true amount of physical activity and inactivity performed. National physical activity studies conducted in the United States and Canada have required a minimum of 10 h·d^{-1} of wear time (Colley et al., 2011; Troiano et al., 2007). Using this 10 h·d^{-1} minimum wear time duration, yeilded 18 to 24 min·d^{-1} of moderate intensity activity for Canadians (Colley et al., 2011) and 18 to 33 min·d^{-1} for US adults (Troiano et al., 2007). Even though the mean accelerometer wear times were approximately 14 h·d^{-1}, data with fewer hours of wear are included in their studies may have attenuated the activity and inactivity values. Our results showed that allowing physical activity data with 10 h·d^{-1} of wear time may be missing 25% to 30% of time spent in inactivity, light, and moderate intensity activity compared to 14 h·d^{-1} of wear time. This could result in inaccurate estimates of the proportion of adults who meet national guidelines of 30 min·d^{-1} or 150 min·wk^{-1} of moderate intensity activity and also underestimate the time spent in inactive behaviors.

Recently, Colley et al. (2010) investigated a variety of accelerometer data reduction methods. One method was comparing 6-, 8-, 10-, 12- and 14- h·d^{-1} of wear time. While only looking at the percentage of the sample that would be included or excluded from analysis with these wear time criteria, the author's noted that lowering the wear time criteria from 14 h·d^{-1} to 10 h·d^{-1} resulted in a substantial increase in valid days whereas lowering from 10 h·d^{-1} to 6 h·d^{-1} only minimally affected the number of valid days for analysis. Their study did not examine the error in minutes per day of activity and inactivity associated with these different wear time values, which we found increased with shorter wear time periods. Colley et al. (2010) showed that accepting a shorter wear time criteria, such as 10 h·d^{-1}, increases the sample size for analysis. However, doing this also increases the error associated with estimates of time spent in physical activity and inactivity and may introduce a sizeable amount of underestimation in the results. For example, our results indicate that when compared to 14 h·d^{-1}, individuals with 10 h·d^{-1} of wear time may be missing roughly 135 min·d^{-1} of inactivity, 95 min·d^{-1} of light, 7 min·d^{-1} of moderate, and 0.5 min·d^{-1} of vigorous activity.

Jerome et al. (2009) examined the difference in a 6 h·d^{-1} vs. a 10 h·d^{-1} of accelerometer wear time criteria and the effect on reliability in a sample of obese individuals. They found similar estimates of time spent in moderate-to-vigorous intensity physical activity between 6 and 10 h·d^{-1} and consequently suggest 6 h·d^{-1} of wear time is adequate in identifying valid days (Jerome, Young, Laferriere, Chen, & Vollmer, 2009). These researchers then used the Spearman-Brown prophecy formula to predict the number of days need to measure to reach an acceptable level of reliability. Using their suggested wear time criteria of 6 h·d^{-1}, the ICC values indicated 16 - 35 days of monitoring would be necessary to achieve a reliability of 0.80 for measuring time spent in moderate-to-vigorous physical activity (Jerome et al., 2009). This amount of monitoring would significantly increase the burden on research participants and increase the cost of conducting a study. In comparison, a study designed to identify sources of variance in daily physical activity, Matthews et al. (2002) analyzed data with more than 12 $h{\cdot}d^{-1}$ of wear time and found that only 3-4 days was necessary to obtain a reliability of ICC > 0.80 (Matthews, Ainsworth, Thompson, et al., 2002). It is unclear if this same result would be found if the wear time criteria were reduced to 10 $h{\cdot}d^{-1}$ of wear time. On the other hand, requiring more than 12 $h{\cdot}d^{-1}$ of wear time as suggested from our results to minimize error, it may be possibly to further reduce the number of days needed to monitor by improving reliability.

If a longer wear time is superior for a more accurately assessment of time spent in inactivity and physical activity intensities, then is it reasonable to have a greater focus on increasing accelerometer wear compliance to ensure that the devices are worn during *all* waking hours or encourage 24 hour wear. A few researchers have had success using 24 $h{\cdot}d^{-1}$ of accelerometer wear time, only suggesting removal for bathing and water activities (Ancoli-Israel et al., 2003; Hofferth, Welk, Treuth, et al., 2008; Sadeh & Acebo, 2002). Little is known if such a practice would yield better measures of time spent in inactivity or in physical activities.

An interesting finding from this study is that the proportion of time spent in inactivity (~54.5%), light (~42.2%), moderate (~3.3%), and vigorous (~0.1%) remained relatively the same across all hours of accelerometer wear time. This shows that our analysis provided little bias by intensity category across wear time. As a result, this information may be helpful for researchers interested in normalizing or adjusting physical activity data based on wear time. Some researchers currently use a form of normalization during data management, commonly normalizing data to 12 hours (Alzahrani et al., 2011; Young et al., 2009). However, data to support the accuracy of this current methodology is limited.

Limitations

This study illustrates the difference in activity estimates if an entire sample has exactly 10-, 11-, 12-, 13-, or 14- h·d^{-1} of accelerometer wear time. Therefore, the absolute differences in minutes of physical activity intensities and inactivity for an entire sample may not be as large because participants are included with a range of accelerometer wear time, even though the minimum criteria may be as low as 10 h·d^{-1}. Nonetheless, this study demonstrates that individuals with less wear time may be adding to an underestimation of the true amount of physical activity that is actually being performed.

Future Research

Since the current study did not stratify data by age, sex, or race/ethnicity, future research should investigate the influence of wear time in different sectors of the population. For example, older adults or highly activity individuals may have different activity patterns that could require different wear time criteria.

Summary

These data illustrate the effect of accelerometer wear time on physical activity data. Allowing data with less wear time may significantly reduce estimates of time spent in activity and inactivity and potentially affect estimates of individuals meeting activity recommendations. This may also influence the results of physical activity interventions which might show individuals were not successful in changing their physical activity or inactivity behaviors when in fact they merely had insufficient accelerometer wear time needed to be accurately assessed and detect significant differences. Sacrificing quality of physical activity assessment by reducing wear time criteria to achieve a greater quantity of participants for analysis may adversely impact study results. This study supports longer accelerometer wear time recommendations of at least 12 $h \cdot d^{-1}$ to ensure accurate estimates of daily physical activity.

References

Alzahrani, M. A., Ada, L., & Dean, C. M. (2011). Duration of physical activity is normal but frequency is reduced after stroke: an observational study. *Journal of Physiotherapy*, *57*.

Ancoli-Israel, S., Cole, R., Alessi, C., Chambers, M., Moorcroft, W., & Pollak, C. (2003). The role of actigraphy in the study of sleep and circadian rhythms. American Academy of Sleep Medicine Review Paper. *Sleep*, *26*(3), 342–92.

Banda, J. A., Hutto, B., Feeney, A., Pfeiffer, K. A., McIver, K., LaMonte, M. J., Blair, S. N., et al. (2010). Comparing physical activity measures in a diverse group of midlife and older adults. *Medicine and Science in Sports and Exercise*, *42*(12), 2251-2257.

Baranowski, T., Baranowski, J., Thompson, D., Buday, R., Jago, R., Griffith, M. J., Islam, N., et al. (2011). Video Game Play, Child Diet, and Physical Activity Behavior Change: A Randomized Clinical Trial. *American Journal of Preventive Medicine*, *40*(1), 33-38.

Boarnet, M. G., Forsyth, A., Day, K., & Oakes, J. M. (2011). The Street Level Built Environment and Physical Activity and Walking: Results of a Predictive Validity Study for the Irvine Minnesota Inventory. *Environment and Behavior*. (published online ahead of print)

Catellier, D. J., Hannan, P. J., Murray, D. M., Addy, C. L., Conway, T. L., Yang, S., & Rice, J. C. (2005). Imputation of missing data when measuring physical activity by accelerometry. *Medicine and Science in Sports and Exercise*, *37*(11 Suppl), S555-562.

Chinapaw, M., Slootmaker, S., Schuit, A., van Zuidam, M., & van Mechelen, W. (2009). Reliability and validity of the Activity

Questionnaire for Adults and Adolescents (AQuAA). *BMC Medical Research Methodology*, *9*(1), 58.

Coleman, K. J., & Epstein, L. H. (1998). Application of generalizability theory to measurement of activity in males who are not regularly active: a preliminary report. *Research Quarterly for Exercise and Sport*, *69*(1), 58-63.

Colley, R. C., Garriguet, D., Janssen, I., Craig, C., Clarke, J., & Tremblay, M. S. (2011). Physical activity of Canadian adults: Accelerometer results from the 2007–2009 Canadian Health Measures Survey. *Health Rep*, *22*(1), 4–11.

Colley, R., Gorber, S. C., & Tremblay, Mark S. (2010). Quality control and data reduction procedures for accelerometry-derived measures of physical activity. *Health Reports / Statistics Canada, Canadian Centre for Health Information = Rapports Sur La Santé / Statistique Canada, Centre Canadien D'information Sur La Santé*, *21*(1), 63-69.

Ekelund, U., Yngve, A., Sjöström, M., & Westerterp, K. (2000). Field evaluation of the Computer Science and Application's Inc. Activity monitor during running and skating training in adolescent athletes. *International Journal of Sports Medicine*, *21*(8), 586-592.

Evenson, K. R. (2011). Towards an understanding of change in physical activity from pregnancy through postpartum. *Psychology of Sport and Exercise*, *12*(1), 36-45.

Evenson, K. R., & Terry, J. W. (2009). Assessment of Differing Definitions of Accelerometer Nonwear Time. *Research Quarterly for Exercise and Sport*, *80*(2), 355-362.

Forsyth, A., Oakes, M. J., Lee, B., & Schmitz, K. H. (2009). The built environment, walking, and physical activity: Is the environment more important to some people than others? *Transportation Research Part D: Transport and Environment*, *14*(1), 42-49.

Freedson, P. S., Melanson, E., & Sirard, J. (1998). Calibration of the Computer Science and Applications, Inc. accelerometer. *Medicine and Science in Sports and Exercise*, *30*(5), 777-781.

Gretebeck, R. J., & Montoye, H. J. (1992). Variability of some objective measures of physical activity. *Medicine and Science in Sports and Exercise*, *24*(10), 1167-1172.

Haskell, W. L., Lee, I. M., Pate, R. R., Powell, K. E., Blair, S. N., Franklin, B. A., Macera, C. 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. *Medicine and Science in Sports and Exercise*, *39*(8), 1423-1434.

Herrmann, S. D., Barriera, T., Kang, M., Ainsworth, B.A. (in review) How Many Hours are Enough? Optimal Accelerometer Wear Time to Reflect Daily Physical Activity.

Hofferth, S. L., Welk, G. J., Treuth, M. S., Randolph, S. M., Curtin, S. C., & Valliant, R. (2008). Validation of a diary measure of children's physical activities. *Sociological Methodology*, *38*(1), 133-154.

Janz, K. F., Witt, J., & Mahoney, L. T. (1995). The stability of children's physical activity as measured by accelerometry and self-report. *Medicine and Science in Sports and Exercise*, *27*(9), 1326-1332.

Jerome, G. J., Young, D. R., Laferriere, Dan, Chen, C., & Vollmer, W. M. (2009). Reliability of RT3 accelerometers among overweight and obese adults. *Medicine and Science in Sports and Exercise*, *41*(1), 110-114.

Kang, M., Bassett, D. R., Barreira, T. V., Tudor-Locke, C., Ainsworth, B. E., Reis, J. P., Strath, S., et al. (2009). How many days are enough? A study of 365 days of pedometer monitoring. *Research Quarterly for Exercise and Sport*, *80*(3), 445-453.

Kang, M., Rowe, D. A., Barreira, T. V., Robinson, T. S., & Mahar, M. T. (2009). Individual information-centered approach for handling physical activity missing data. *Research Quarterly for Exercise and Sport*, *80*(2), 131-137.

Kang, M., Zhu, W., Tudor-Locke, C. & Ainsworth, B. (2005). Experimental Determination of Effectiveness of an Individual Information-Centered Approach in Recovering Step-Count Missing Data. *Measurement in Physical Education and Exercise Science*, *9*(4), 233-250.

Levin, S., Jacobs, D. R., Ainsworth, B. E., Richardson, M. T., & Leon, A. S. (1999). Intra-individual variation and estimates of usual physical activity. *Annals of Epidemiology*, *9*(8), 481-488.

Mâsse, L. C., Fuemmeler, B. F., Anderson, C. B., Matthews, C. E., Trost, S. G., Catellier, D. J., & Treuth, M. (2005). Accelerometer data reduction: a comparison of four reduction algorithms on select outcome variables. *Medicine and Science in Sports and Exercise*, *37*(11 Suppl), S544-554.

Macera, C. A., Ham, S. A., Yore, M. M., Jones, D. A., Ainsworth, B. E., Kimsey, C. D., & Kohl, H. W. (2005). Prevalence of physical activity in the United States: Behavioral Risk Factor Surveillance System, 2001. *Preventing Chronic Disease*, *2*(2), A17.

Matthews, C. E., Chen, K. Y., Freedson, P. S., Buchowski, M. S., Beech, B. M., Pate, R. R., & Troiano, R. P. (2008). Amount of time spent in sedentary behaviors in the United States, 2003–2004. *American journal of epidemiology*, *167*(7), 875-881.

Matthews, C. E., Ainsworth, B. E., Thompson, R. W., & Bassett, D. R. (2002). Sources of variance in daily physical activity levels as measured by an accelerometer. *Medicine and Science in Sports and Exercise*, *34*(8), 1376-1381.

Motl, R. W., Dlugonski, D., Suh, Y., Weikert, M., Fernhall, B., & Goldman, M. (2010). Accelerometry and Its Association With Objective Markers of Walking Limitations in Ambulatory Adults With Multiple

Sclerosis. *Archives of Physical Medicine and Rehabilitation*, *91*(12), 1942-1947.

Murray, D. M., Catellier, D. J., Hannan, P. J., Treuth, M. S., Stevens, J., Schmitz, K. H., Rice, J. C., et al. (2004). School-level intraclass correlation for physical activity in adolescent girls. *Medicine and Science in Sports and Exercise*, *36*(5), 876-882.

Pate, R. R., Pratt, M., Blair, S. N., Haskell, W. L., Macera, C. A., Bouchard, C., Buchner, D., et al. (1995). Physical activity and public health. A recommendation from the Centers for Disease Control and Prevention and the American College of Sports Medicine. *JAMA: The Journal of the American Medical Association*, *273*(5), 402-407.

Sadeh, A., & Acebo, C. (2002). The role of actigraphy in sleep medicine. *Sleep Medicine Reviews*, *6*(2), 113–124.

Semanik, P., Song, J., Chang, R. W., Manheim, L., Ainsworth, B., & Dunlop, D. (2010). Assessing physical activity in persons with rheumatoid arthritis using accelerometry. *Medicine and Science in Sports and Exercise*, *42*(8), 1493-1501.

Slootmaker, S. M., Schuit, A. J., Chinapaw, M. J., Seidell, J. C., & van Mechelen, W. (2009). Disagreement in physical activity assessed by accelerometer and self-report in subgroups of age, gender, education and weight status. *The International Journal of Behavioral Nutrition and Physical Activity*, *6*, 17.

Spittaels, H., Verloigne, M., Gidlow, C., Gloanec, J., Titze, S., Foster, C., Oppert, J. M., et al., others. (2010). Measuring physical activity-related environmental factors: reliability and predictive validity of the European environmental questionnaire ALPHA. *International Journal of Behavioral Nutrition and Physical Activity*, *7*(1), 48.

Swartz, A. M., Strath, S. J., Bassett, D. R., O'Brien, W. L., King, G. A., & Ainsworth, B. E. (2000). Estimation of energy expenditure using CSA accelerometers at hip and wrist sites. *Medicine and Science in Sports and Exercise*, *32*(9 Suppl), S450-456.

Treuth, M. S., Sherwood, N. E., Butte, N. F., McClanahan, B., Obarzanek, E., Zhou, A., Ayers, C., et al. (2003). Validity and reliability of activity measures in African-American girls for GEMS. *Medicine and Science in Sports and Exercise*, *35*(3), 532-539.

Troiano, R. P., Berrigan, D., Dodd, K. W., MâSse, L. C., Tilert, T., & Mcdowell, M. (2007). Physical Activity in the United States Measured by Accelerometer. *Medicine & Science in Sports & Exercise*, *40*(1):181-188

Trost, S. G., Pate, R. R, Freedson, P. S., Sallis, J. F., & Taylor, W. C. (2000). Using objective physical activity measures with youth: how many days of monitoring are needed? *Medicine & Science in Sports & Exercise*, *32*(2), 426-431.

Tudor-Locke, C., Burkett, L., Reis, J. P., Ainsworth, B. E., Macera, C. A., & Wilson, D. K. (2005). How many days of pedometer monitoring predict weekly physical activity in adults? *Preventive Medicine*, *40*(3), 293-298.

Young, D. R., Jerome, G. J., Chen, C., Laferriere, D., & Vollmer, W. M. (2009). Patterns of physical activity among overweight and obese adults. *Preventing Chronic Disease*, *6*(3), A90.

Chapter 9

DISCUSSION

Physical activity and inactivity are global health concerns that demand rigorous research and effective population interventions to increase the number of people that are sufficiently active to receive health benefits from physical activity. This dissertation compiled five projects that highlighted and evaluated a variety of methods use in physical activity measurement. Understanding the measurement capabilities of instruments and improving current methodologies is essential to the accurate assessment and monitoring of physical activity.

Project one evaluated a new commercial accelerometer (MyWellness Key). The MyWellness key was designed for use by the general public to accurately monitor their physical activity level. The goal of this device is to allow individuals to accurately assess and increase their awareness of their physical activity level. The individual can then use the device to track their physical activity to maintain or increase to levels that are encouraged to obtain health benefits. The purpose of this project was to compare the Mywellness Key to other criterion physical activity measures. The results demonstrated that the MyWellness Key compared favorably with these measures and has the potential to be a valuable commercially available tool for physical activity monitoring. This type of validity evidence is crucial to understand prior to widespread implementation of an instrument.

Project two was designed evaluate a physical activity self-report questionnaire (GPAQ) that is currently being used by the WHO for international physical activity surveillance. The GPAQ has been used in over 50 countries to assess physical activity levels. The intention is to allow for inter and intra country comparisons. However, to date, there is only limited validity and reliability evidence of the GPAQ to support its use for physical activity assessment. This project investigated evidence of validity and 10 day test-retest reliability of the GPAQ. The results showed fair-to-moderate evidence of validity and high test-retest reliability. These results are similar to other studies which have examined subjective physical activity questionnaires that are currently recommended for use in physical activity surveillance. Physical activity surveys offer value because they can be administered widely with a much lower cost that most objective physical activity instruments. While some measurement validity is sacrificed due to limitations of self-report questionnaires (e.g., recall bias, social desirability, etc.), this type of assessment tool can provide useful information to researchers and policy makers.

Project three was devised to address the lack of consensus on data reduction criteria, specifically wear time, used in accelerometer studies of physical activity. Currently, there are wide ranging criteria for deciding whether or not a day of physical activity monitoring by accelerometry represents a valid day. This project compared a range of wear time criteria (10 hours per day to 14 hours per day) to identify the optimal wear time needed to assess daily physical activity at a variety of activity intensity levels. The results showed that allowing less than 13 hours per day of wear time can significantly underestimate the true amount of physical performed. Due to these results, the comparability of studies that use different wear time criteria and suggest that many studies may be underestimating physical activity levels must be questioned. More research in this area is needed but this demonstrated a clear need for standardizing wear time criteria among studies.

Project four focused on emerging technology and a new measures (MAPS score) to assess the interactions between a person’s physical activity within their environment using accelerometers, GPS receivers, and GIS. The MAPS score was developed to assess free-living function. However, there is little information about the validity of this score in a healthy population. The results showed the MAPS score has high reliability and modest validity evidence as compared with measures of physical activity and correlates of physical activity. These data indicated that MAPS measures a different construct than physical activity or environmental measures alone and that additional studies are needed to understand the constructs that the MAPS score measures. In comparison with the use of survey measures, using an objective assessment of person-environment interaction displayed in the MAPS score has the potential to enhance measurement capabilities of studies interested in physical activity behavior or the influence of neighborhood environments on physical activity.

Project five examined the effects of different accelerometer wear time criteria in participants from the 2005-2006 NHANES. This study was designed to expand upon project three by using similar methods and analyses in a larger, national representative sample. Different amounts of wear time were compared (10 hours per day to 14 hours per day) to understand the impact of these criteria on physical activity data. The results confirmed the conclusion from project three that allowing less than 13 hours per day of accelerometer wear time may result in a significant underestimation of time spent in various physical activity intensity levels. This underestimation may affect the accuracy of individuals being classified as meeting physical activity guidelines or being classified as insufficiently activity. Understanding the optimal parameters to assess physical activity and inactivity are important steps in promoting a standardized method for identifying and ensuring a valid day of accelerometer assessment.

The projects that comprised this dissertation reiterate the fact that physical activity is a complex behavior that requires a variety of methods to accurately assess physical activity levels and monitor changes. Achieving the public health goal of increasing physical activity to health enhancing levels can be helped by striving to improve understanding of new and current instruments and advance methodologies in physical activity assessment.

REFERENCES

ActiGraph. (2009). *ActiLife Users Manual*. Pensacola, FL: ActiGraph; p. 8-51.

Ainsworth, B. E., Haskell, W. L., Leon, A. S., Jacobs, D. R., Montoye, H. J., Sallis, J. F., & Paffenbarger, R. S. (1993). Compendium of physical activities: classification of energy costs of human physical activities. *Medicine and Science in Sports and Exercise*, *25*(1), 71-80.

Anderson, J. (1971). Space - time budgets and activity studies in urban georaphy and planning. *Environment and Planning*, *3*(4), 353 – 368.

Armstrong, T., & Bull, F. (2006). Development of the World Health Organization Global Physical Activity Questionnaire (GPAQ). *Journal of Public Health*, *14*(2), 66-70.

Aschengrau, A. & Seage, G. R. (2003). *Essentials of Epidemiology in Public Health Practice.* Subbury, MA: Jones and Bartlett Publishers; p. 251-279.

Bassett, D. R., Wyatt, H. R., Thompson, H., Peters, J. C., & Hill, J. O. (2010). Pedometer-measured physical activity and health behaviors in US adults. *Medicine & Science in Sports & Exercise*, *42*(10), 1819.

Bauman, A., Bull, F., Chey, T., Craig, C., Ainsworth, B., Sallis, J., Bowles, H., et al. (2009). The International Prevalence Study on Physical Activity: results from 20 countries. *International Journal of Behavioral Nutrition and Physical Activity*, *6*(1), 21.

Bauman, A., & Craig, C. (2005). The place of physical activity in the WHO Global Strategy on Diet and Physical Activity. *International Journal of Behavioral Nutrition and Physical Activity, 2*(1), 10.

Bauman, A. E., Sallis, J. F. (2008). Global problems require global studies. *American Journal of Preventive Medicine*, *34*(6), 544-545.

Beets, M. W., Bornstein, D., Beighle, A., Cardinal, B. J., & Morgan, C. F. (2010). Pedometer-measured physical activity patterns of youth: a 13-country review. *American Journal of Preventive Medicine*, *38*(2), 208-216.

Blair, S. N., Haskell, W. L., Ho, P., Paffenbarger, R. S., Vranizan, K. M., Farquhar, J. W., & Wood, P. D. (1985). Assessment of habitual physical activity by a seven-day recall in a community survey and controlled experiments. *American Journal of Epidemiology*, *122*(5), 794-804.

Blair, S. N., LaMonte, M. J., & Nichaman, M. Z. (2004). The evolution of physical activity recommendations: how much is enough? *American Journal of Clinical Nutrition*, *79*(5), S913-920.

Boon, R. M., Hamlin, M. J., Steel, G. D., & Ross, J. J. (2010). Validation of the New Zealand Physical Activity Questionnaire (NZPAQ-LF) and the International Physical Activity Questionnaire (IPAQ-LF) with Accelerometry. *British Journal of Sports Medicine, 44*(10)*, 741-746*.

Bouchard, C., Tremblay, A., Leblanc, C., Lortie, G., Savard, R., & Thériault, G. (1983). A method to assess energy expenditure in children and adults. *The American Journal of Clinical Nutrition*, *37*(3), 461-467.

Brage, S, Brage, N., Franks, P. W., Ekelund, U, & Wareham, N J. (2005). Reliability and validity of the combined heart rate and movement sensor Actiheart. *European Journal of Clinical Nutrition*, *59*(4), 561-570.

Brown, W. J., Trost, S. G., Bauman, A., Mummery, K., & Owen, N. (2004). Test-retest reliability of four physical activity measures used in population surveys. *Journal of Science and Medicine in Sport / Sports Medicine Australia*, *7*(2), 205-215.

Bull, F. C., Maslin, T. S., & Armstrong, T. (2009). Global physical activity questionnaire (GPAQ): nine country reliability and validity study. *Journal of Physical Activity and Health*, *6*(6), 790-804.

Caspersen, C. J., Powell, K. E., & Christenson, G. M. (1985). Physical activity, exercise, and physical fitness: definitions and distinctions for health-related research. *Public Health Reports*, *100*(2), 126-131.

Catellier, D. J., Hannan, P. J., Murray, D. M., Addy, C. L., Conway, T. L., Yang, S., & Rice, J. C. (2005). Imputation of missing data when measuring physical activity by accelerometry. *Medicine and Science in Sports and Exercise, 37*, S555-S562.

Centers for Disease Control and Prevention (CDC). National Center for Health Statistics (NCHS). (2010). National Health and Nutrition Examination Survey Examination Protocol. Retrieved December 9, 2010, from http://www.cdc.gov/nchs/nhanes.htm

Chan, C. B., Ryan, D. A. J., & Tudor-Locke, C. (2004). Health benefits of a pedometer-based physical activity intervention in sedentary workers. *Preventive Medicine*, *39*(6), 1215-1222.

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*(1), 155-159.

Craig, C. L., Cameron, C., Griffiths, J. M., & Tudor-Locke, C. (2010). Descriptive epidemiology of youth pedometer-determined physical activity: CANPLAY. *Medicine and Science in Sports and Exercise*, *42*(9), 1639-1643.

Craig, C. L., Marshall, A. L., Sjöström, M., Bauman, A. E., Booth, M. L., Ainsworth, B. E., Pratt, M., et al. (2003). International physical activity questionnaire: 12-country reliability and validity. *Medicine and Science in Sports and Exercise*, *35*(8), 1381-1395.

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

Crouter, S. E., Churilla, J. R., & Bassett, D. R. (2008). Accuracy of the Actiheart for the assessment of energy expenditure in adults. *European Journal of Clinical Nutrition*, *62*(6), 704-711.

Crouter, S. E., Churilla, J. R., & Bassett, D. R. (2006). Estimating energy expenditure using accelerometers. *European Journal of Applied Physiology, 98*, 601-612.

Crouter, S. E., Schneider, P. L., Karabulut, M., & Bassett, D. R. (2003). Validity of 10 electronic pedometers for measuring steps, distance, and energy cost. *Medicine and Science in Sports and Exercise*, *35*(8), 1455-1460.

Duncan, M. J., Mummery, W. K., & Dascombe, B. J. (2007). Utility of global positioning system to measure active transport in urban areas. *Medicine and Science in Sports and Exercise*, *39*(10), 1851-1857.

Ekelund, U., Sepp, H., Brage, S., Becker, W., Jakes, R., Hennings, M., & Wareham, N. J. (2006). Criterion-related validity of the last 7-day, short form of the International Physical Activity Questionnaire in Swedish adults. *Public Health Nutrition*, *9*(2), 258-265.

Evenson, K. R. & Terry, J. W. (2009). Assessment of differing definitions of accelerometer nonwear time. *Research Quarterly for Exercise and Sport*, *80*(2), 355-362.

Fogelholm, M., Malmberg, J., Suni, J., Santtila, M., Kyröläinen, H., Mäntysaari, M., & Oja, P. (2006). International Physical Activity Questionnaire: Validity against fitness. *Medicine and Science in Sports and Exercise*, *38*(4), 753-760.

Frank, L. D., Schmid, T. L., Sallis, J. F., Chapman, J., & Saelens, B. E. (2005). Linking objectively measured physical activity with objectively measured urban form: Findings from SMARTRAQ. *Am J Preventive Medicine, 28*(2), 117-125.

Freedson, P. S., Melanson, E., & Sirard, J. (1998). Calibration of the Computer Science and Applications, Inc. accelerometer. *Medicine and Science in Sports and Exercise*, *30*(5), 777-781.

Godin, G., & Shephard, R. J. (1985). A simple method to assess exercise behavior in the community. *Canadian Journal of Applied Sport Sciences. Journal Canadien Des Sciences Appliquées Au Sport*, *10*(3), 141-146.

Golledge, R. G., & Stimson, R. J. (1997). *Spatial behavior: a geographic perspective*. Guilford Press.

Gregg, E. W., Gerzoff, R. B., Caspersen, C. J., Williamson, D. F., & Narayan, K. M. V. (2003). Relationship of Walking to Mortality Among US Adults With Diabetes. *Archives of Internal Medicine*, *163*(12), 1440-1447.

Guthold, R., Ono, T., Strong, K. L., Chatterji, S., & Morabia, A. (2008). Worldwide Variability in Physical Inactivity: A 51-Country Survey. *American Journal of Preventive Medicine, 34*(6):486-494.

Hagerstrand, T. (1970). What about people in regional science? (Vol. 24, pp. 7-21). Presented at the Regional Science Association.

Haggett, P. (1965). *Locational Analysis in Human Geography.* (1st ed.). Edward Arnold.

Hagströmer, M., Oja, P., & Sjöström, M. (2006). The International Physical Activity Questionnaire (IPAQ): a study of concurrent and construct validity. *Public Health Nutrition*, *9*(6), 755-762.

Hagströmer, M., Oja, P., & Sjöström, M. (2007). Physical activity and inactivity in an adult population assessed by accelerometry. *Medicine and Science in Sports and Exercise, 39*(9), 1502-1508.

Hallal, P. C., Simoes, E., Reichert, F. F., Azevedo, M. R., Ramos, L. R., Pratt, Michael, & Brownson, R. C. (2010). Validity and reliability of the telephone-administered international physical activity questionnaire in Brazil. *Journal of Physical Activity & Health*, *7*(3), 402-409.

Ham, S. A., Reis, J. P., Strath, S. J., Dubose, K. D., & Ainsworth, B. E. (2007). Discrepancies between methods of identifying objectively determined physical activity. *Medicine & Science in Sports & Exercise, 39*, 52-58.

Haskell, W. L., Lee, I. M., Pate, R. R., Powell, K. E., Blair, S. N., Franklin, B. A., Macera, C. 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. *Circulation*, *116*(9), 1081.

Hatano, Y. (1993). Use of the pedometer for promoting daily walking exercise. *International Council for Health, Physical Education, and Recreation*, *29*(4), 4–8.

Hendelman, D., Miller, K., Baggett, C., Debold, E., & Freedson, P. (2000) Validity of accelerometry for the assessment of moderate intensity physical activity in the field. *Medicine and Science in Sports and Exercise, 32*, 442–449.

Herrmann, S. D., Barriera, T., Kang, M., Ainsworth, B.A. (in review) How Many Hours are Enough? Optimal Accelerometer Wear Time to Reflect Daily Physical Activity.

Herrmann, S. D., Hart, T. L., Lee, C. D., & Ainsworth, B. E. (2011). Evaluation of the MyWellness Key accelerometer. *British Journal of Sports Medicine, 45*:109-113.

Herrmann, S. D., & Ragan, B. G. (2008). Outcome Assessment Part 2: Measurement of Person-Environment Interaction. *Athletic Therapy Today*, *13*(2), 24-27.

Herrmann, S. D., Ragan, B. G., Scott, C. B., Mack, M. G., Kang, M., & Dompier, T. P. (2008). Development and Validation of a Movement and Activity in Physical Space (MAPS) Score as a Functional Outcome Measure. *Journal of Athletic Training*, *43*(2), S.

Herrmann, S. D., Snook, E. M., Kang, M., Scott, C. B., Mack, M. G., Dompier, T. P., & Ragan, B. G. (in press). Development and Validation of a Movement and Activity in Physical Space (MAPS) Score as a Functional Outcome Measure. *Archives of Physical Medicine and Rehabilitation*.

Hendelman, D., Miller, K., Baggett, C., Debold, E., Freedson, P. (2000). Validity of accelerometry for the assessment of moderate intensity physical activity in the field. . *Medicine and Science in Sports and Exercise*, *32*(9):S442-449.

Jacobs, D. R., Ainsworth, B. E., Hartman, T. J., & Leon, A. S. (1993). A simultaneous evaluation of 10 commonly used physical activity questionnaires. *Medicine and Science in Sports and Exercise*, *25*(1), 81-91.

Jia, Y., Xu, L., Kang, D., & Tang, Y. (2008). Reliability and validity regarding the Chinese version of the International Physical Activity Questionnaires (long self-administrated format) on women in Chengdu, China. *Zhonghua Liu Xing Bing Xue Za Zhi*, *29*(11), 1078-1082.

Kang, M., Bassett, D. R., Barreira, T. V., Tudor-Locke, C., Ainsworth, B. E., Reis, J. P., Strath, S., et al. (2009). How Many Days Are Enough? A Study of 365 Days of Pedometer Monitoring. *Research Quarterly for Exercise and Sport*, *80*(3), 445-453.

Kang, M., Rowe, D. A., Barreira, T. V., Robinson, T. S., & Mahar, M. T. (2009). Individual information-centered approach for handling physical activity missing data. *Research Quarterly for Exercise and Sport*, *80*(2):131–137.

Kang, M., Zhu, W., Tudor-Locke, C., & Ainsworth, B. (2005). Experimental determination of effectiveness of an individual information-centered approach in recovering step-count missing data. *Measurement in Physical Education and Exercise Science*, *9*, 233–250.

Kayes, N. M., Schulter, P. J., McPherson, K. M., Leete, M., Mawston, G., & Taylor, D. (2009). Exploring Actical accelerometers as an objective measure of physical activity in people with multiple sclerosis. *Archives of Physical Medicine & Rehabilitation, 90*, 594-601.

Keyserling T. C., Hodge C. D., Jilcott S. B., Johnston, L. F., Garcia, B. A., Gizlice, Z., et al. (2008). Randomized trial of a clinic-based, community-supported, lifestyle intervention to improve physical activity and diet: The North Carolina enhanced WISEWOMAN project. *Preventive Medicine, 46,* 499-510.

Klesges, R. C., Eck, L. H., Mellon, M. W., Fulliton, W., Somes, G. W., & Hanson, C. L. (1990). The accuracy of self-reports of physical activity. *Medicine and Science in Sports and Exercise, 22*(5), 690-697.

Kraus, W. E., Houmard, J. A., Duscha, B. D., Knetzger, K. J., Wharton, M. B., McCartney, J. S., Bales, C. W., et al. (2002). Effects of the amount and intensity of exercise on plasma lipoproteins. *The New England Journal of Medicine*, *347*(19), 1483-1492.

Lagerros, Y. T., & Lagiou, P. (2007). Assessment of physical activity and energy expenditure in epidemiological research of chronic diseases. *European Journal of Epidemiology*, *22*(6), 353-362.

LaMonte, M. J., Blair, S. N., & Church, T. S. (2005). Physical activity and diabetes prevention. *Journal of Applied Physiology*, *99*(3), 1205-1213.

LaMonte, M. J., & Ainsworth, B. E. (2001). Quantifying energy expenditure and physical activity in the context of dose response. *Medicine & Science in Sports & Exercise*, *33*(6), S370-378.

LaMonte, M. J., Ainsworth, B. E., & Reis, J. P. (2006). Measuring Physical Activity. In T. M. Wood & W. Zhu (Eds.), *Measurement Theory and Practice in Kinesiology*. Champaign IL: Human Kinetics.

LaPorte, R. E., Adams, L. L., Savage, D. D., Brenes, G., Dearwater, S., & Cook, T. (1984). The spectrum of physical activity, cardiovascular disease and health: an epidemiologic perspective. *American Journal of Epidemiology*, *120*(4), 507-517.

Le Faucheur, A., Abraham, P., Jaquinandi, V., Bouyé, P., Saumet, J. L., & Noury-Desvaux, B. (2007). Study of human outdoor walking with a low-cost GPS and simple spreadsheet analysis. *Medicine and Science in Sports and Exercise*, *39*(9), 1570-1578.

Le Faucheur, A., Abraham, P., Jaquinandi, V., Bouyé, P., Saumet, J. L., & Noury-Desvaux, B. (2008). Measurement of walking distance and speed in patients with peripheral arterial disease: a novel method using a global positioning system. *Circulation*, *117*(7), 897-904.

Lee, I.-M., Sesso, H. D., Oguma, Y., & Paffenbarger, Ralph S. (2004). The "Weekend Warrior" and Risk of Mortality. *American Journal of Epidemiology*, *160*(7), 636 -641.

Lenntorp, B. (1976). *Paths in space-time environments a time-geographic study of movement possibilities of individuals*. Lund: Gleerup.

Li, T. Y., Rana, J. S., Manson, J. E., Willett, W. C., Stampfer, M. J., Colditz, G. A., Rexrode, K. M., et al. (2006). Obesity as Compared With Physical Activity in Predicting Risk of Coronary Heart Disease in Women. *Circulation*, *113*(4), 499-506.

Macera, C. A., Ham, S. A., Yore, M. M., Jones, D. A., Ainsworth, B. E., Kimsey, C. D., & Kohl, H. W. (2005). Prevalence of physical activity in the United States: behavioral risk factor surveillance system, 2001. *Preventing Chronic Disease*, *2*(2), 1-10.

Macfarlane, D. J., Lee, C. C. Y., Ho, E. Y. K., Chan, K. L., & Chan, D. T. S. (2007). Reliability and validity of the Chinese version of IPAQ (short, last 7 days). *Journal of Science and Medicine in Sport*, *10*(1), 45-51.

Makikawa, M. & Murakami, D. (1996). Development of an ambulatory physical activity and behavior map monitoring system. Presented at the International Conference of the IEEE Engineering in Medicine and Biologic Society, Amsterdam, Netherlands.

Matthews, C. E. (2005). Calibration of accelerometer output for adults. *Medicine and Science in Sports and Exercise, 37*, S512-522.

Matthews, C. E., Freedson, P. S., Hebert, J. R., Stanek, E. J., Merriam, P. A., Rosal, M. C., Ebbeling, C. B., et al. (2001). Seasonal variation in household, occupational, and leisure time physical activity: longitudinal analyses from the seasonal variation of blood cholesterol study. *American Journal of Epidemiology*, *153*(2), 172-183.

Matthews, C. E., Hebert, J. R., Freedson, P. S., Stanek, E. J., Merriam, P. A., Ebbeling, C. B., & Ockene, I. S. (2001). Sources of variance in daily physical activity levels in the seasonal variation of blood cholesterol study. *American Journal of Epidemiology*, *153*(10), 987-995.

Matthews, C. E. (2002). Use of Self-Report Instruments to Assess Physical Activity. In G. J. Welk (Ed.), *Physical Activity Assessments for Health-Related Research*. Champaign IL: Human Kinetics.

Matthews, C. E., Ainsworth, B. E., Thompson, R. W., & Bassett, D. R. (2002). Sources of variance in daily physical activity levels as measured by an accelerometer. *Medicine and Science in Sports and Exercise*, *34*(8), 1376-1381.

Matthews, C. E., Ainsworth, B. E., Hanby, C., Pate, R. R., Addy, C., Freedson, P. S., Jones, D. A., et al. (2005). Development and testing of a short physical activity recall questionnaire. *Medicine and Science in Sports and Exercise*, *37*(6), 986-994.

Matthews, C. E., Chen, K. Y., Freedson, Patty S., Buchowski, M. S., Beech, B. M., Pate, R. R., & Troiano, R. P. (2008). Amount of Time Spent in Sedentary Behaviors in the United States, 2003–2004. *American Journal of Epidemiology*, *167*(7), 875-881.

McClain, J., Abraham, T., Brusseau, T., & Tudor-Locke, C. (2008). Epoch Length and Accelerometer Outputs in Children: Comparison to Direct Observation. *Medicine and Science in Sports and Exercise, 40*(12), 2080-2087.

Murphy, S. L. (2009). Review of physical activity measurement using accelerometers in older adults: Considerations for research design and conduct. *Preventive Medicine, 48*, 108-114.

Mokdad, A. H., Marks, J. S., Stroup, D. F., & Gerberding, J. L. (2004). Actual Causes of Death in the United States, 2000. *Journal of the American Medical Association*, *291*(10), 1238-1245.

Montoye, H. J., Kemper, H. C. G., Saris, W. H. M., & Washburn, R. A. (1996). *Measuring physical activity and energy expenditure*. Champaign IL: Human Kinetics.

Morinder, G., Mattsson, E., Sollander, C., Marcus, C., & Larsson, U. E. (2009). Six-minute walk test in obese children and adolescents: Reproducibility and validity. *Physiotherapy Research International*, *14*(2), 91-104.

Murphy, M., Nevill, A., Neville, C., Biddle, S., & Hardman, A. (2002). Accumulating brisk walking for fitness, cardiovascular risk, and psychological health. *Medicine and Science in Sports and Exercise*, *34*(9), 1468-1474.

Oja, P., Bull, F., Fogelholm, M., & Martin, B. (2010). Physical activity recommendations for health: what should Europe do? *BMC Public Health*, *10*(1), 10.

Padilla, J., Wallace, J. P., & Park, S. (2005). Accumulation of physical activity reduces blood pressure in pre-and hypertension. *Medicine & Science in Sports & Exercise*, *37*(8), 1264.

Paffenbarger Jr, R. S., Hyde, R., Wing, A. L., & Hsieh, C. (1986). Physical activity, all-cause mortality, and longevity of college alumni. *New England journal of medicine*, *314*(10), 605–613.

Paffenbarger, R. S., Wing, A. L., & Hyde, R. T. (1978). Physical activity as an index of heart attack risk in college alumni. *American Journal of Epidemiology*, *108*(3), 161-175.

Paffenbarger, R. S., Kampert, J. B., & Lee, I. M. (1997). Physical activity and health of college men: longitudinal observations. *International Journal of Sports Medicine*, *18 Suppl 3*, S200-203.

Palm, R. (1981). Women in nonmetropolitan areas: a time-budget survey. *Environment and Planning A*, *13*(3), 373-378.

Papathanasiou, G., Georgoudis, G., Georgakopoulos, D., Katsouras, C., Kalfakakou, V., & Evangelou, A. (2010). Criterion-related validity of the short International Physical Activity Questionnaire against exercise capacity in young adults. *European Journal of Cardiovascular Prevention and Rehabilitation*, *17*(4), 380-386.

Pate, R. R., Pratt, M., Blair, S. N., Haskell, W. L., Macera, C. A., Bouchard, C., Buchner, D., et al. (1995). Physical activity and public health: A recommendation from the Centers for Disease Control and Prevention and the American College of Sports Medicine. *JAMA*, *273*(5), 402-407.

Patrick, K., Griswold, W. G., Raab, F., & Intille, S. S. (2008). Health and the mobile phone. *American Journal of Preventive Medicine*, *35*(2), 177-181.

Popkin, B. M., Kim, S., Rusev, E. R., Du, S., & Zizza, C. (2006). Measuring the full economic costs of diet, physical activity and obesity-related chronic diseases. *Obesity Reviews*, *7*(3), 271–293.

Pratt, M., Epping, J. N., & Dietz, W. H. (2009). Putting physical activity into public health: a historical perspective from the CDC. *Preventive Medicine*, *49*(4), 301-302.

Quigg, R., Gray, A., Reeder, A. I., Holt, A., & Waters, D. L. (2010). Using accelerometers and GPS units to identify the proportion of daily physical activity located in parks with playgrounds in New Zealand children. *Preventive Medicine*, *50*(5-6), 235-240.

Rodríguez, D. A., Brown, A. L., & Troped, P. J. (2005). Portable global positioning units to complement accelerometry-based physical activity monitors. *Medicine and Science in Sports and Exercise*, *37*(11 Suppl), S572-581.

Rowe, D. A., & Maher, M. T. (2006). Validity. In T. M. Wood & W. Zhu (Eds.), *Measurement Theory and Practice in Kinesiology*. Champaign, IL: Human Kinetics.

Rütten, A., Ziemainz, H., Schena, F., Stahl, T., Stiggelbout, M., Vanden Auweele, Y., et al. (2003). Using different physical activity measurements in eight European countries. Results of the European Physical Activity Surveillance System (EUPASS) time series survey. *Public Health Nutrition, 6*(4):371-376.

Rzewnicki, R., Vanden Auweele, Y., & De Bourdeaudhuij, I. (2003). Addressing overreporting on the International Physical Activity Questionnaire (IPAQ) telephone survey with a population sample. *Public Health Nutrition*, *6*(3), 299-305.

Saarloos, D., Kim, J. E., & Timmermans, H. (2009). The Built Environment and Health: Introducing Individual Space-Time Behavior, *6*(6), 1724-1743.

Sadeh, A. & Acebo, C. (2002). The role of actigraphy in sleep medicine. *Sleep Medicine Reviews, 6*(2), 113-124.

Sallis, J. F. & Saelens, B. E. (2000). Assessment of physical activity by self-report: status, limitations, and future directions. *Research Quarterly for Exercise & Sport, 71*(2):S1-14.

Schmidt, M. D., Freedson, P. S., & Chasan-Taber L. (2003). Estimating physical activity using the CSA accelerometer and a physical activity log. *Medicine & Science in Sports & Exercise, 35*, 1605-1611.

Schneider, P. L., Crouter, S. E., & Bassett, D. R. (2004). Pedometer measures of free-living physical activity: comparison of 13 models. *Medicine & Science in Sports & Exercise*, *36*(2), 331-335.

Schneider, P. L., Crouter, S. E., Lukajic, O., & Bassett, D. R. (2003). Accuracy and reliability of 10 pedometers for measuring steps over a 400-m walk. *Medicine & Science in Sports & Exercise*, *35*(10), 1779-1784.

Schönfelder, S., & Axhausen, K. W. (2002). Measuring the size and structure of human activity spaces: The longitudinal perspective. *Arbeitsbericht Verkehrs-und Raumplanung*, *135*.

Schönfelder, S., Axhausen, K., Antille, N., & Bierlaire, M. (2002). Exploring the potentials of automatically collected GPS data for travel behaviour analysis - a Swedish data source. In Möltgen, J. and Wytzisk, A., (Ed.), GI-Technologien für Verkehr und Logistik, 13: 155–179.

Schutz, Y., & Herren, R. (2000). Assessment of speed of human locomotion using a differential satellite global positioning system. *Medicine and Science in Sports and Exercise*, *32*(3), 642-646.

Sharpe, P. A., Wilcox, S., Rooney, L. J., Strong, D., Hopkins-Campbell, R., & Butel, J. (In Press). Adherence to accelerometer protocols among women from economically disadvantaged neighborhoods. *Journal of Physical Activity and Health*.

Shephard, R., & Vuillemin, A. (2003). Limits to the measurement of habitual physical activity by questionnaires. *British Journal of Sports Medicine*, *37*(3), 197-206.

Slattery, M. L., & Jacobs, D. R. (1995). Assessment of ability to recall physical activity of several years ago. *Annals of Epidemiology*, *5*(4), 292-296.

Slootmaker, S. M., Schuit, A. J., Chinapaw, M. J., Seidell, J. C., & van Mechelen, W. (2009). Disagreement in physical activity assessed by accelerometer and self-report in subgroups of age, gender, education and weight status. *International Journal of Behavioral Nutrition and Physical Activity, 6*(1), 17.

Snook, E. M., Scott, C. B., Ragan, B. G., Morand, A. M., & Sackau, T. (2010). Combining GPS and Accelerometer Data as a New Functional Outcome Measure for Multiple Sclerosis: A Case Comparison Study of the Movement and Activity in Physical Space [MAPS] Score. Presented at the mHealth Summit Conference, Washington, DC.

Spittaels, H., Verloigne, M., Gidlow, C., Gloanec, J., Titze, S., Foster, C., Oppert, J. M., et al., others. (2010). Measuring physical activity-related environmental factors: reliability and predictive validity of the European environmental questionnaire ALPHA. *International Journal of Behavioral Nutrition and Physical Activity*, *7*(1), 48.

Swartz, A. M., Strath, S. J., Bassett, D. R., O'Brien, W. L., King, G. A., Ainsworth, B. E. (2000). Estimation of energy expenditure using CSA accelerometers at hip and wrist sites. *Medicine & Science in Sports & Exercise, 32*, S450-456.

Tapia, E., Intille, S., Haskell, W., Larson, K., Wright, J., King, A., & Friedman, R. (2007). Real-Time Recognition of Physical Activities and Their Intensities Using Wireless Accelerometers and a Heart Rate Monitor. *Proceedings of the 2007 11th IEEE International Symposium on Wearable Computers* (pp. 1-4). Boston, MA, USA: IEEE Computer Society.

Taylor, H. L., Jacobs, D. R., Schucker, B., Knudsen, J., Leon, A. S., & Debacker, G. (1978). A questionnaire for the assessment of leisure time physical activities. *Journal of Chronic Diseases*, *31*(12), 741-755.

Thomas, S., Reading, J., & Shephard, R. J. (1992). Revision of the physical activity readiness questionnaire (PAR-Q). *Canadian Journal of Sports Science, 17*(4)*,* 338-345.

Trinh, O. T. H., Nguyen, N. D., van der Ploeg, H. P., Dibley, M. J., & Bauman, A. (2009). Test-retest repeatability and relative validity of the Global Physical Activity Questionnaire in a developing country context. *Journal of Physical Activity and Health*, *6(*1), S46-53.

Troiano, R. P., Berrigan, D., Dodd, K. W., Mâsse, L. C., Tilert, T., & Mcdowell, M. (2007). Physical Activity in the United States Measured by Accelerometer. *Medicine & Science in Sports & Exercise*, *40*(1):181-188.

Troped, P. J., Oliveira, M. S., Matthews, C. E., Cromley, E. K., Melly, S. J., & Craig, B. A. (2008). Prediction of Activity Mode with Global Positioning System and Accelerometer Data. *Medicine & Science in Sports & Exercise*, *40*(5), 972-978.

Troped, P. J., Wilson, J. S., Matthews, C. E., Cromley, E. K., & Melly, S. J. (2010). The Built Environment and Location-Based Physical Activity. *American Journal of Preventive Medicine*, *38*(4), 429-438.

Trost, S. G., Pate, R. R., Freedson, P. S., Sallis, J. F., & Taylor, W. C. (2000). Using objective physical activity measures with youth: how many days of monitoring are needed? *Medicine and Science in Sports and Exercise*, *32*(2), 426-431.

Tudor-Locke, C., & Bassett, D. R. (2004). How many steps/day are enough? Preliminary pedometer indices for public health. *Sports Medicine (Auckland, N.Z.)*, *34*(1), 1-8.

Tudor-Locke, C., Brashear, M. M., Johnson, W. D., & Katzmarzyk, P. T. (2010). Accelerometer profiles of physical activity and inactivity in normal weight, overweight, and obese U.S. men and women. *The International Journal of Behavioral Nutrition and Physical Activity*, *7*, 60.

Tudor-Locke, C., Burkett, L., Reis, J. P., Ainsworth, B. E., Macera, C. A., & Wilson, D. K. (2005). How many days of pedometer monitoring predict weekly physical activity in adults? *Preventive Medicine*, *40*(3), 293-298.

Tudor-Locke, C., Hatano, Y., Pangrazi, R. P., & Kang, M. (2008). Revisiting “how many steps are enough?” *Medicine and Science in Sports and Exercise*, *40*(7 Suppl), S537-543.

Tudor-Locke, C., McClain, J. J., Hart, T. L., Sisson, S. B., & Washington, T. L. (2009). Pedometry Methods for Assessing Free-living Youth. *Research Quarterly for Exercise & Sport, 80*(2), 175-184.

Tudor-Locke, C., Sisson, S. B., Collova, T., Lee, S. M., & Swan, P. D. (2005). Pedometer-determined step count guidelines for classifying walking intensity in a young ostensibly healthy population. *Canadian Journal of Applied Physiology = Revue Canadienne De Physiologie Appliquée*, *30*(6), 666-676.

Tudor-Locke, C., Williams, J. E., Reis, J. P., & Pluto, D. (2002). Utility of pedometers for assessing physical activity: convergent validity. *Sports Medicine*, *32*(12), 795-808.

US Department of Health and Human Services (USDHHS). (1996). *Physical activity and health: a report of the Surgeon General*. National Center for Chronic Disease Prevention and Health Promotion. Atlanta, Georgia: US

US Department of Health and Human Services (USDHHS). (2008). *Physical Activity Guidelines Advisory Committee Report*. Washington, DC: USDHHS; p. vi-14.

World Health Organization | Physical activity. (n.d.). Retrieved October 6, 2010, from http://www.who.int/dietphysicalactivity/publications/facts/pa/en/index.html

Wannamethee, S. G., Shaper, A. G., & Walker, M. (1998). Changes in physical activity, mortality, and incidence of coronary heart disease in older men. *The Lancet*, *351*(9116), 1603-1608.

Warburton, D. E., Nicol, C. W., & Bredin, S. S. (2006). Health benefits of physical activity: the evidence. *CMAJ*, *174*(6), 801-809.

Ward, D. S., Evenson, K. R., Vaughn, A., Rodgers, A. B., & Troiano, R. P. (2005). Accelerometer use in physical activity: best practices and research recommendations. *Medicine and Science in Sports and Exercise, 37*(11), S582.

Washburn, R. A., Heath, G. W., & Jackson, A. W. (2000). Reliability and validity issues concerning large-scale surveillance of physical activity. *Research Quarterly for Exercise and Sport*, *71*(2 Suppl), S104-113.

Welk, G. J. (2002). *Physical Activity Assessments for Health-Related Research*. Champaign, IL: Human Kinetics.

Welk, G. J. (2005) Principles of design and analyses for the calibration of accelerometry-based activity monitors. *Medicine and Science in Sports and Exercise, 37*, S501-11.

Welk, G. J., Blair, S. N., Wood, K., Jones, S., & Thompson, R. W. (2000). A comparative evaluation of three accelerometry-based physical activity monitors. *Medicine and Science in Sports and Exercise*, *32*, S489-97.

Wilde, B. E., Sidman, C. L., & Corbin, C. B. (2001). A 10,000-step count as a physical activity target for sedentary women. *Research Quarterly for Exercise and Sport*, *72*(4), 411-414.

Wolf, J., Guensler, R., & Bachman, W. (2001). Elimination of the Travel Diary: Experiment to Derive Trip Purpose from Global Positioning System Travel Data. *Transportation Research Record: Journal of the Transportation Research Board*, *1768*(1), 125-134.

Wood, Terry M., & Zhu, Weimo. (2006). *Measurement theory and practice in kinesiology*. Human Kinetics.

World Health Organization. (2002). The World Health Report 2002: Reducing risks, promoting healthy lifestyles. Geneva, Switzerland: World Health Organization: 248.

Young, D. R., Jerome, G. J., Chen, C., Laferriere, D., Vollmer, W. M. (2009). Patterns of physical activity among overweight and obese adults. *Preventing Chronic Disease, 6*, 1-9.

Zhu, W. (2003). Assessing physical activity space: Issues and challenges. *Proceedings II of 2003 Daegu Universiade Conference: Facing the Challenge*, 2003 Daegu Universiade Conference Organizing Committee, 601-608.