

1-1-2017

Predicting cardiovascular fitness in ethnic minority youth: A comparison of demographic, body composition, and physical activity variables

Steven Arcidiacono

Nova Southeastern University, sjarcidi@gmail.com

This document is a product of extensive research conducted at the Nova Southeastern University [College of Psychology](#). For more information on research and degree programs at the NSU College of Psychology, please [click here](#).

Follow this and additional works at: http://nsuworks.nova.edu/cps_stueta

 Part of the [Psychology Commons](#)

Share Feedback About This Item

NSUWorks Citation

Arcidiacono, S. (2017). Predicting cardiovascular fitness in ethnic minority youth: A comparison of demographic, body composition, and physical activity variables. .

Available at: http://nsuworks.nova.edu/cps_stueta/112

This Dissertation is brought to you by the College of Psychology at NSUWorks. It has been accepted for inclusion in College of Psychology Theses and Dissertations by an authorized administrator of NSUWorks. For more information, please contact nsuworks@nova.edu.

**PREDICTING CARDIOVASCULAR FITNESS IN ETHNIC MINORITY
YOUTH: A COMPARISON OF DEMOGRAPHIC, BODY COMPOSITION,
AND PHYSICAL ACTIVITY VARIABLES**

by

Steve Arcidiacono, M.S.

A Dissertation Presented to the College of Psychology
of Nova Southeastern University
in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy

NOVA SOUTHEASTERN UNIVERSITY

2016

DISSERTATION APPROVAL SHEET

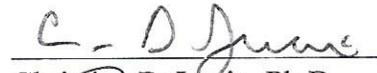
This dissertation was submitted by Steven J. Arcidiacono under the direction of the Chairperson of the dissertation listed below. It was submitted to the School of Psychology and approved in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Clinical Psychology at Nova Southeastern University.

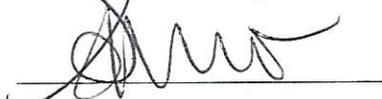
08/11/2016
Date of Defense

11/30/2016
Date of Final Approval

Approved:


David Reitman, Ph.D., Chair


Christian DeLucia, Ph.D.


Stephen A. Russo, Ph.D.


David Reitman, Ph.D., Chair

Table of Contents

LIST OF TABLES	vii
LIST OF FIGURES	ix
ABSTRACT	x
CHAPTER I: STATEMENT OF THE PROBLEM	1
Introduction	1
Methodological review of Body Composition and Physical Activity Measures ...	7
Body Composition Measures	11
Air Displacement Plethysmography (ADP).....	12
Dual-Energy X-ray Absorptiometry (DXA).....	13
Hydrodensitometry	14
Body Mass Index (BMI)	15
Skinfold Thickness.....	17
Bioelectrical Impedance Analysis (BIA)	19
Body composition measurement summary	22
Measuring Youth Physical Activity	24
Behavioral Observation	25
Doubly Labeled Water (DLW)	30
Pedometer	31
Accelerometry.....	33
Physical activity measurement conclusions.....	38
Clinical utility of body composition and physical activity	41
The “gold standard” problem and multiple instrument utilization	42
The Relationship between Body Composition, Physical Activity, and Physical Fitness	43
The role of demographic variables	48

PREDICTING YOUTH FITNESS

CHAPTER II: METHOD	51
Participants.....	51
Measures	52
Body Mass Index (BMI)	52
Bioelectrical Impedance Analysis (BIA)	53
ActiGraph GT1M Accelerometer	54
Progressive Aerobic Cardiovascular Endurance Run (PACER) and Healthy Fitness Zone (HFZ)	56
Procedures.....	57
Analyses	59
Hypotheses	59
Statistical Methods.....	60
Accelerometry Data	60
Statistical Analysis.....	61
CHAPTER III: RESULTS	63
Descriptive Statistics.....	63
Correlational Analysis	67
Regression Analyses	67
CHAPTER IV: DISCUSSION	72
Implications of Findings	72
Interpretation of Findings by Analysis	74
BMI and BIA-measured body fat	74
Physical activity and ethnic/gender subgroups	75
Correlational analysis.....	76
Regression analysis.....	77
Limitations	79
Future Research	83
REFERENCES	86

PREDICTING YOUTH FITNESS

ACKNOWLEDGMENTS

I would first like to thank my committee for their relentless support, guidance, and commitment to my progress on a clinical, professional, and personal level. I am grateful for my chair, Dr. David Reitman, for his critical feedback, patience, commitment to rigorous research, and for pushing me to persevere through the numerous barriers that I faced, including complicated data analyses, expired software licenses, personal stressors, and my own anxiety. Dr. Christian DeLucia has always been a second mentor for me, and I thank him for teaching me to be skeptical in how I read and interpret research, as well as his thoughtful advice and expertise in the methodological and statistical aspects of this project. I appreciate Dr. Russo for his clinical guidance, shrewd advice, and helping me to choose a topic that merges my passion in the areas of health, technology, and interventions with children and adolescents. I would like to thank Dr. Jean Thaw for her essential feedback at critical stages of this project, and for allowing me to utilize data gathered as part of her grant-funded research. This dissertation would not be possible without support from the Active Living Research Program of the Robert Wood Johnson Foundation, and by The Children's Trust. I would also like to thank the other members of the Project Rise team for their time, energy, effort, and fidelity in collecting this data; notably, Drs. Manuela Villa, Leon Mandler, and Yalemni Luna DeLaurier.

I would also like to thank my clinical supervisors, intern cohort, and colleagues of Boys Town in Nebraska and South Florida for their support, recommendations, pep talks, patience and encouragement throughout this process. To Drs. Danielle Overton, Lydia Malcolm, and Jessica Ketterer, thank you for boosting

PREDICTING YOUTH FITNESS

me up and showing me how to conquer adversity. I am ever grateful to my parents, family, and friends for supporting me in every facet and helping me keep my life balanced. I must acknowledge my best friend and wife, Jessica, for her love, patience, and understanding throughout this entire process. She supported me when I needed support, pushed me when I needed to be pushed, and took care of life tasks when I needed to focus. Finally, I would like to dedicate this dissertation to the memory of Erin Conant, whose friendship made it possible for me to persevere when I moved a thousand miles from home, and to my grandfather, Tom Stead, a lifelong teacher whose lessons about hard work, growth, and success still guide me today.

PREDICTING YOUTH FITNESS

LIST OF TABLES

Table 1: BMI Prevalence \geq 95th Percentile in Children and Adolescents age 6-19

(Ogden et al., 2010)

Table 2: Validity, Reliability and Clinical Utility Criteria of Youth Body Composition Measures

Table 3: Validity, Reliability, and Clinical Utility of Physical Activity Measurements

Table 4: Hierarchical Regression Results Predicting Fitness in 8th Grade Girls

(Lohman, 2008)

Table 5: Descriptive Statistics of Sample

Table 6: Descriptive Characteristics

Table 7: Percentage of Time Spent in Physical Activity Categories on Weekdays and Weekends

Table 8: Daily Minutes in Physical Activity Categories

Table 9: Percent of Validated Time in Physical Activity Categories

Table 10: Two-Way Analysis of Variance of Physical Activity by Gender and Race/Ethnicity

Table 11: Bivariate Correlations

Table 12: Hierarchical Regression Variables Contributing to PACER Laps

PREDICTING YOUTH FITNESS

Table 13: Hierarchical Regression Variables Contributing to PACER Laps (Age Excluded)

Table 14: HFZ Descriptive Statistics

Table 15: Logistic Regression Variables Predicting HFZ Categorization

Table 16: Logistic Regression Variables Predicting HFZ Categorization (Age and Gender Excluded)

LIST OF FIGURES

Figure 1: Psychometric properties and clinical utility of youth body composition measures

Figure 2: Psychometric properties and overall clinical utility of physical activity measures in children and adolescents

Figure 3: Percent of time engaging in different intensities of physical activity by ethnic-gender subgroup

**PREDICTING CARDIOVASCULAR FITNESS IN ETHNIC MINORITY
YOUTH: A COMPARISON OF DEMOGRAPHIC, BODY COMPOSITION,
AND PHYSICAL ACTIVITY VARIABLES**

by

Steve Arcidiacono, M.S.

Nova Southeastern University

ABSTRACT

Prevalence of obesity, low physical activity, and poor physical fitness of youth in the United States are increasingly poor and in need of intervention to prevent later concerns like hypertension. The overall goal of this dissertation was to examine which factors weigh heaviest in predicting cardiovascular fitness in diverse youth, and how we might measure those factors by maximizing clinical utility and psychometric properties. The sample was gathered from a larger study examining physical activity in youth from Miami-Dade county enrolled in out-of-school programs. Participants ($N = 58$) were aged 6-17 and comprised exclusively of Hispanic and Non-Hispanic Black children and adolescents, the majority of whom were from low-income families. Predictors of fitness were gathered in three primary categories: demographic variables (age, gender, race/ethnic category, family income level), body composition (Body Mass Index [BMI] percentile, Bioelectrical Impedance Analysis [BIA]-measured body fat percentage), and habitual physical

activity (accelerometer-measured counts per minute). These factors were entered in a hierarchical regression model to predict cardiorespiratory fitness measured by performance on a 20-meter shuttle run. Physical activity was not found to be significantly associated with fitness, and the effect size of this relationship was small, particularly when considering the impact of demographic and body composition variables. Overall, results reinforced the need for interventions to improve body composition and increase physical activity: the average participant was at the 81st percentile of BMI, had 26% body fat, was sedentary for approximately 84% of awake time, and only spent a few minutes per day engaging in vigorous physical activity. There were significant main effects of gender and race/ethnic category such that males and Non-Hispanic Black participants generally spent a greater proportion of time engaging in physical activity, with less sedentary time. Being female, younger, and having less body fat was associated with performance in the healthy fitness range when considering the impact of other variables, even though boys and older participants had more laps on the shuttle run. Findings presented in this dissertation indicate a continued need to develop technology with high utility, validity, and reliability to measure and improve indicators of health in diverse, low-income youth.

CHAPTER I: STATEMENT OF THE PROBLEM

Introduction

Obesity rates in childhood and adolescence have risen to alarming levels over the last 30 years and continue to be a considerable health concern in the United States (Hedley et al., 2004; Ogden, Yanovski, Carroll, & Flegal, 2007). Recent research indicates that nearly one fifth of U.S. children are obese, as defined by a body mass index (BMI) at or above the 95th percentile of expected BMI for age (Ogden, Carrol, Curtin, Lamb, & Flegal, 2010). Over 30% of U.S. youth could be considered overweight as defined by a BMI at or above the 85th percentile of expected BMI for age (Ogden et al., 2010). Moreover, the prevalence of youth obesity is higher in Hispanic and Black ethnic groups compared to Non-Hispanic White youth (Ogden et al., 2010; see Table 1). These rates vary by gender such that for males, the highest proportion of obesity is among Hispanic youth, followed by Black youth, and White youth. In contrast, the highest prevalence of obesity for females is Black youth, followed by Hispanic, and White youth (see Table 1).

Table 1

BMI Prevalence \geq 95th Percentile in Children and Adolescents age 6-19 (Ogden et al., 2010)

	Overall	Hispanic	Non-Hispanic White	Non-Hispanic Black
Boys	20.1	26.7	18.2	18.9
Girls	17.3	19.5	15.6	25.9
All	18.7	23.2	17	22.3

Being obese or overweight during childhood may result in negative psychosocial consequences and has been linked to diseases such as asthma, early onset diabetes, and hypertension (Wabitsch, 2000a, 2000b). In 2005, Kirk et al.

performed a treatment outcome study aimed at reducing BMI and determining associated health outcomes using 117 obese youth aged 5 to 19. Researchers found an overall reduction in BMI using a treatment involving group exercise, parent education, and behavioral intervention therapy to improve diet and physical activity over a mean of approximately 6 months. This improvement in BMI was also associated with improved total cholesterol, low-density lipoprotein (LDL) cholesterol, triglycerides, and insulin. While lack of a control group and high attrition rate (55%) limit conclusions that can be drawn from the effectiveness of the treatment, it is important to note that reductions in BMI are associated with positive changes to important indicators of health in children and adolescents.

While BMI is a practical, commonly used indicator of health, limitations of BMI (see BMI section below) led some to suggest that more direct and accurate measures of body fat should also be considered in estimating obesity (Prentice & Jebb, 2001). In 2011, Going et al. investigated the connection between percent body fat and chronic disease risk factors in U.S. children and adolescents using data from the National Health and Nutritional Examination Surveys (NHANES) III (1998-1994) and IV (1999-2004). This large-scale longitudinal study included data from over 12,000 diverse children and adolescents (36% Black, 36% Mexican-American, 28% White) aged 6-18. Results indicated that increased body fat as measured by skin calipers in childhood (e.g., above 20% for boys and 30% for girls) is associated with higher prevalence of unhealthy levels of cardiovascular disease risk factors such as systolic blood pressure, HDL cholesterol, insulin and glucose (Going et al., 2011). Furthermore, obese children are at a greater risk of becoming obese as adults, which

has been linked with an increased risk of serious health issues such as cardiovascular disease and type 2 diabetes (Ogden et al., 2007; Serdula et al., 1993). Such research suggests a need for prevention and intervention efforts to counter the negative consequences of the obesity epidemic present in the United States.

The cause of obesity has been identified by the Centers for Disease Control and Prevention (CDC, 2009) as an imbalance involving excessive calorie consumption and/or inadequate physical activity. Furthermore, it is suggested that for each individual, body weight is a combination of genetic, metabolic, behavioral, environmental, cultural, and socioeconomic influences. One study investigating trends in caloric intake indicated that relatively steady rates of caloric intake from 1910 to 1985 increased by roughly 300 calories per day between 1985 and 2000 (Putnum, Allshouse, & Kantor, 2002). While 300 calories per day may not initially seem to be a substantial amount (equivalent to about one large soft drink at a fast food restaurant), this totals to over 100,000 calories (equivalent to about 30lbs of body weight) added per year. Since the causes of obesity appear to be multifaceted, prevention and intervention efforts will likely need to address the multiple issues with diet and exercise as primary concerns (CDC, 2009; Finkelstein, Ruhm, & Kosa, 2005).

Recent strategies identified by the CDC (2009) aimed at encouraging physical activity and limiting sedentary activity among children and youth include increasing the amount of physical activity in physical education programs and increasing opportunities for extracurricular physical activity. In 2004, Patrick et al. investigated diet, physical activity, and sedentary behaviors as potential risk factors for being

overweight (BMI \geq 85th percentile) in a diverse sample (42% ethnic minority) of 878 adolescents. Results indicated that for both boys and girls, being overweight was associated with less vigorous activity time, less fiber consumption, and, interestingly, less energy consumed. Some explanation for seemingly contradictory finding is that the increased vigorous activity time in the normal group may lead to greater energy expenditure and thus, necessitate the need for increased energy consumption. For boys, television viewing on nonschool days was also associated with overweight status. Using a multivariate logistical regression model, Patrick et al. (2004) found that after accounting for age, ethnicity, education level, and total energy intake, insufficient vigorous activity was the only significant risk factor for elevated BMI. Other dietary factors such as fat and fiber consumption were not significant risk factors in this sample. Overall, such results indicate that low physical activity is a key risk factor of being overweight in adolescence.

One systematic review found that adolescent physical activity itself may have numerous healthy consequences (Hallal, Victora, Azevedo, & Wells, 2006). Here, physical activity was found to be associated with both short-term benefits such as lowered systolic and diastolic blood pressure, and long-term benefits such as improved bone health and lowered risk for breast cancer (Hallal et al., 2006). Adolescent physical activity also seemed to improve aspects of mental health such as less depressive symptoms and improved self-esteem (Hallal et al., 2006). Furthermore, longitudinal research indicates that BMI and physical activity during adolescence is strongly related to adult BMI, supporting the notion that physical activity at a young age leads to physical activity in adulthood (Kvaavik, Tell, &

Klepp, 2003). Given this evidence, physical activity is likely a key component to improving both the short term and long term health of children and adolescents.

While it is plausible that body composition and physical activity are related, empirical support for this relationship is inconsistent, due to variance in methodology and measurement (Abbott & Davies, 2004; Rennie et al., 2005). Furthermore, physical activity and physical fitness are sometimes used interchangeably in colloquial use; however, in the current study an important distinction will be maintained. Physical activity has been defined by Ortega et al. (2008) as, “any body movement produced by muscle action that increases energy expenditure” (p. 2). By contrast, physical fitness is the capacity for an individual to perform physical activity and includes reference to other qualities such as cardiorespiratory capacity. Thus, physical fitness by definition includes physical activity but physical activity does not necessarily involve other indicators of physical fitness. Research indicates that physical fitness and physical activity are likely to be closely related to one another but findings are limited by the complexity of assessing physical activity, particularly in children and adolescents (Kohl, Fulton, & Capersen, 2000; Ortega et al., 2008). Other findings also indicate that low physical activity and poor physical fitness are direct and independent predictors of negative health consequences often attributed to obesity, such as type 2 diabetes and metabolic dysfunction (Telford, 2007).

Physical fitness is often indicated by cardiorespiratory fitness or the capacity of the cardiovascular system to perform during prolonged exercise. Widely considered to be the best indicator of cardiovascular fitness is the maximal oxygen consumption attained during a graded exercise to voluntary exhaustion (VO_2^{max} ;

Shephard et al., 1968). The most common test for assessing youth cardiorespiratory fitness in epidemiological studies has been the 20-meter shuttle run (e.g., Progressive Aerobic Cardiovascular Endurance Run [PACER]; Ortega et al., 2008).

Cardiovascular fitness has been found to be an important indicator of health as it has been associated with body fatness measured by skinfold thickness as well as dual-energy X-ray absorptiometry (see below for descriptions; Gutin, Yin, Humphries, & Barbeau, 2005; Ortega, et al., 2007; Ruiz et al., 2006). Longitudinal findings also indicate that fitness in adolescence is associated with adulthood fitness and adiposity (Eisenmann, Wickel, Welk, & Blair, 2005). Furthermore, physical fitness in childhood and adolescence has been associated with important cardiovascular risk factors and metabolic profiles (e.g., triglycerides, cholesterol, insulin; see Ortega et al., 2008). In addition, according to Ortega et al. (2008), there is empirical evidence that suggests, “the deleterious consequences ascribed to high fatness could be counteracted by having high levels of cardiorespiratory fitness” (p.4). Such findings suggest that interventions to improve health should aim not just to reduce weight but also improve cardiovascular fitness. There is some evidence that suggests that high-intensity physical activity is related to improvements in physical fitness (see below; Gutin et al., 2005; Ruiz et al., 2006). However, the relationship between such variables in children and adolescence is complex and requires further research; primarily due to the complexity of the relationship and difficulty in measuring complex variables such as physical activity (Ortega et al., 2008). Notwithstanding the influence of diet and other epidemiological factors; the current study aims to clarify the relationship between body composition, physical activity, and

cardiovascular fitness while considering the influence of important demographic factors such as age, gender, and ethnicity.

Methodological Review of Body Composition and Physical Activity Measures

As previously discussed, some of the difficulty in drawing conclusions involving the relationship between physical activity, body composition, and physical fitness is due to methodological differences and lack of agreement on a “gold standard” measure of body composition or physical activity. In order to understand the results of studies utilizing different instrumentation as well as the strengths and weaknesses of each measure, the current proposal first evaluates several major instruments on the basis of their psychometric properties and clinical utility. Following this review, relevant studies utilizing multiple instruments to investigate the relationships between body composition, physical activity, and physical fitness will be discussed to represent the present knowledge and lingering questions to be addressed by the current study.

The current study aims to first evaluate body composition and physical activity measurement in children and adolescents with consideration of recent movement toward evidence-based assessment (EBA; Kazdin, 2005; Mash & Hunsley, 2005). This review aims to evaluate measures with consideration of important purposes of assessment, identified by Mash and Hunsley (2005), including: (a) diagnosis and case formulation (i.e., determining the causes of negative health status; categorization of youth into categories such as underweight, overweight, or obese) (b) screening (i.e., identifying those that are, or are at risk for being, overweight and obese) (c) prognosis (i.e., potential health outcomes that can result from overweight

and obesity along with benefits that can result from increased physical activity) (d) treatment design and planning (i.e., developing interventions to improve body composition and physical activity) (e) treatment monitoring (i.e., tracking changes in body composition, physical activity, and other relevant variables such as health status) (d) treatment evaluation (i.e., the effectiveness, ability to be applied across contexts and populations, and cost effectiveness of the intervention). Kazdin (2005) argued that an assessment must include adequate and appropriate psychometric qualities (i.e., reliability, validity) as well as other criteria (e.g., utility, user friendliness) in order for a measure to meet a given purpose.

Reliability has been defined as the extent to which data are consistent (Huck, 2008). This concept can be addressed using multiple indices including internal consistency (i.e., parts of a measure contribute in a consistent way and measure the same thing), interrater reliability (i.e., consistency in results if used or scored by different raters), and test-retest reliability (i.e., stability of results if the measure is completed a second time; Huck, 2008; Mash & Hunsley, 2005). Several common statistics used to represent different types of reliability include internal consistency coefficients (split-half correlation; Cronbach's alpha), indices of interrater reliability (Pearson's r , Cohen's kappa, intraclass correlation [ICC]) retest correlations (Pearson's r ; ICC), and typical error of measurement (standard error of measurement [SEM]; coefficient of variation [CV]; Huck, 2008). These methods of measuring consistency are distinct and a high score on one measure of reliability does not ensure that the instrument will have a high score on another measure of reliability (Huck, 2008). According to Kazdin (2005), "there are many types... not all are needed,

important, or relevant for a given purpose or measure.” Sechrest, Stickle, and Stewart (1998) indicate that reliability estimates often are not directly relevant to the purpose of the instrument. For example, test-retest reliability may be very important for an instrument used in monitoring change over time (e.g., evaluating the impact of a physical activity intervention during the school year) but less meaningful when the instrument is to be used to screen for obesity. Thus, it is essential to consider the appropriateness of a method of estimating reliability and do so in the context of the concept being measured.

Campbell’s validity typology (Campbell, 1957; Campbell & Stanley, 1963) introduced the concepts of internal and external validity, and outlined several potential threats to validity. In general, Campbell’s concept of internal validity asks to what extent the observed association between variables signifies a causal effect. In the context of intervention research, internal validity asks to what extent the intervention, and not another factor, caused the outcome. Three criteria are generally utilized to establish causality: temporal precedence (cause precedes effect), statistical association (quantified relationship), and nonspuriousness (ruled out confounding factors). According to Campbell, several threats to internal validity should be considered, including selection (systematic pretreatment differences), history (events that occur during treatment), and maturation (naturally occurring changes that could account for treatment effect). One should ensure that there are no systematic differences between groups at baseline as results may be due to these pre-treatment differences rather than the treatment itself.

There are several types of validity that represent different ways to assess accuracy. One common representation of instrument validity is construct validity, which asks which components of the intervention are responsible for the observed difference or change. This includes the extent to which the instrument exhibits convergent validity (i.e., the instrument is associated with instruments that measure a similar concept) as well as divergent validity (i.e., the instrument is not associated with instruments that measure other constructs). Another important element of instrument validity is criterion-related validity (i.e., the instrument's association with a relevant outcome), which includes two subtypes: data collected at the same time (concurrent validity), and outcomes measured in the future (predictive validity; Huck, 2008). Sechrest et al. (1998) refer to the combination of reliability and validity as "enlightenment." While enlightenment is undoubtedly an essential aspect of an instrument, recent perspectives in EBA indicate that enlightenment is not sufficient in determining the appropriateness of a measure in a given context.

In addition to demonstrating reliability and validity (i.e., enlightenment) there are several additional criteria to assess clinical utility, which are important to consider when using an instrument to monitor progress over time (see Kazdin, 2005).

Instruments are needed that are: (a) acceptable to both the person being evaluated and test administrators (e.g., reasonable, relevant, and worthwhile) (b) feasible given common restraints (e.g., cost-effective, user-friendly) (c) bidirectional (e.g., there is an ability to measure both gains and losses) (d) able to retain validity over time (e.g., resistant to practice effects, reactivity) and (e) meaningful in degrees or level of change (e.g., easily understood in real-world contexts). While "enlightenment" is a

highly laudable goal, consideration of the less-commonly discussed criteria above is critical to understanding an instrument's clinical utility.

There has been difficulty in establishing a consensus about ideal methodological approaches and measurement tools for measures of body composition and physical activity used with children and adolescents. Measures of obesity vary considerably in their reliability and utility, leading to difficulty in comparing results across studies (Wheeler & Twist, 2010). Youth movement often occurs in erratic patterns that are difficult to predict and measure, such as rapid, brief, bursts of moderate to vigorous physical activity (Bailey et al., 1995; Hands, Parker, & Larkin, 2006). Measures of body composition and physical activity also have limitations when considered in light of the clinical utility criteria suggested by Kazdin (e.g., lack of feasibility due to high cost, lack of portability, invasiveness or burden on participants; lack of meaningfulness in degree of change due to ambiguous data or the need for extensive training to interpret results). Therefore, the aim of this section is to critically evaluate current measures of body composition and physical activity in children and adolescents via consideration of each measure's psychometric adequacy (i.e. enlightenment) and clinical utility.

Body composition measures. Measurement of youth body composition is challenging due to physiological changes that occur throughout development (Lohman, 1986); thus, there are a number of methods to measure body composition in children and adolescents. The instruments reviewed in the current study represent the most prominently studied measures based on searches for youth body composition-related studies in several databases (e.g., ProQuest, SAGE, ScienceDirect, Wiley).

Measures typically utilized in laboratory settings are presented first (ADP, DXA, Hydrodensitometry), followed by measures typically used in field research (BMI, skinfold thickness, BIA). Each instrument is reviewed systematically with regard to: enlightenment (reliability and validity) followed by clinical utility (acceptability, feasibility, resistance to reactivity, bidirectionality, and meaningfulness of data).

Air Displacement Plethysmography (ADP). ADP is a procedure in which whole body volume is calculated by determining the change in volume of a chamber with and without the participant inside. Bone density, weight and other biometric values are then used to determine percent body fat estimates. The reliability of percent body fat measured by ADP in children is often underreported but reliability estimates include test-retest reliability on the same day ($r = .98$) and multiple-day coefficient of variation ($CV = 5.3\%$; Anderson, 2007). ADP has been empirically supported as a fast and accurate measure of body composition in obese and non-obese populations (Fields, Goran, & McCrory, 2002; Ginde et al., 2005).

Acceptability of ADP in children is supported by the relative ease of administration, but limited by the potential discomfort of participants being enclosed in the chamber (Fields et al., 2002; Ginde et al., 2005). The primary practical advantages of ADP include the ability for the unit to quickly assess body composition and accommodate large participants (up to 500lbs; Ginde et al., 2005). However, feasibility is limited because units are not easily transported from one location to another and may cost upwards of \$30,000-40,000 (Fields et al., 2002). ADP has advantages in bidirectionality as body fat percentage gains and losses can easily be observed. Furthermore, ADP does not provide immediate feedback to participants,

reducing the likelihood of participant reactivity. A further strength of ADP is the meaningfulness in degrees of change in outcomes, since body fat in weight or percentage can be easily understood as a relevant indicator of obesity. Overall, the psychometric properties of ADP and other characteristics (e.g. ease of administration, accommodation of large participants) show promise; however, feasibility constraints such as high cost and lack of portability limit its clinical utility.

Dual-Energy X-ray Absorptiometry (DXA). DXA uses x-ray beams of varying energy levels in order to establish bone density, soft tissue, and (through subtraction) fat content. Empirical findings indicate a high rate of consistency in measuring percent body fat in children (+/- 2-4%; Ellis, Shypailo, Pratt, & Pond, 1994). However, concerns about radiation exposure sometimes limit repeated measurements and thus, test-retest reliability, in children (e.g., Lazzer et al., 2008). DXA has been validated in its association with (yet distinction from) a number of similar instruments such as ADP and hydrodensitometry (see below; Lazzer et al., 2008).

Regarding the utility criteria as suggested by Kazdin (2005), DXA has evidenced acceptability in research contexts as it has been used in a number of studies as the criterion method to which other body composition methods are compared (Elberg et al., 2004; Mooney et al., 2011). Feasibility is improved by the short duration of administration but limited by relatively expensive equipment costs, lack of portability, and required expert training in order to accurately administer and interpret results in children and adolescents (Bachrach, 2000; Fields et al., 2002; Tyrrell et al., 2001). Body fat percentage as an outcome is bidirectional and easily

understood in real world contexts. Ethical considerations regarding exposure to radiation in children may limit the extent to which the instrument may be used to track change over time (Lazzer et al., 2008). Overall, the demonstrated enlightenment and acceptability of DXA as a criterion measure remain strengths but limitations in feasibility and utility in measuring change over time limit DXA as a useful tool in certain contexts.

Hydrodensitometry. Hydrodensitometry, also known as underwater weighing, involves the measurement of body volume before and while the participant is submerged in water. Equations regarding bone and muscle density are then used to calculate body composition. For example, a person with more bone and muscle will have a greater weight while submerged, indicating greater body density and, subsequently, a lower percent of body fat. Reliability estimates of hydrodensitometry are limited but have been found to be adequate and better in children compared to adults (coefficients of variation between 6-10%; Demerath et al., 2002). Hydrodensitometry has been found to be similar, but distinct (i.e., demonstrating construct validity), in determining body fat when compared to ADP and DXA (Demerath et al., 2002; Fields et al., 2002).

Regarding clinical utility, hydrodensitometry has a number of limitations. Acceptability is limited as the instrument requires participants to fully exhaust the air in their lungs while completely submerged underwater. Unfortunately, data indicates that between 13-17% of children are unable or unwilling to complete the requirements of hydrodensitometry (Harsha & Bray, 1996; Reilly, Wilson, & Durnin, 1995). Additionally, feasibility is limited as the procedure requires specialized

training and facilities (Demerath et al., 2002). Again, the bidirectionality of percent body fat is an advantage as gains and losses are easily understood. Another strength is the lack of participant reactivity due to lack of participant feedback. While body fat appears to be meaningful in the “real-world,” the method used to determine body fat is less direct (i.e., face-valid) compared to other measures (e.g., DXA) and instead relies on assumptions about hydration and body composition and uses equations to make estimates (Tyrell et al., 2001). As a whole, hydrodensitometry seems adequate psychometrically, but is severely limited in its clinical utility in children, primarily due to the low level of clinical acceptability associated with submerging children underwater.

Body Mass Index (BMI). BMI is one of the most commonly used methods to assess whether a person is overweight or obese (Prentice & Jebb, 2001). In order to calculate the BMI, one divides the person’s weight in kilograms by the square of their height in meters. This value is then compared to age and gender-specific cutoffs in order to determine the person’s comparative health. In children and adolescents, there are four weight status categories: Underweight (Less than 5th percentile), Healthy Weight (5th percentile to less than 85th percentile), Overweight (85th to less than 95th percentile), and Obese (Equal to or greater than the 95th percentile; CDC, 2011). It is suggested that weight and height be measured by investigators using scales and stadiometers (Himes, 2009). Self-report as an alternative way of obtaining height and weight measurements may be less burdensome and time consuming than using electric scales and stadiometers; yet, empirical evidence suggests that in youth, height is often overestimated and weight is often underestimated, resulting in biased

BMI values that are 2-3 units (kg/m²) lower than BMI values obtained through biometric measures (Sherry, Jefferds, & Grummer-Strawn, 2007). Since the range of healthy BMI in childhood and adolescence is typically between 5-10 units and the range between overweight and obese is typically between 2-5 units, this error may be crucial in underestimating overweight and obese cases. When measurements are obtained correctly, BMI has been supported as a reliable and valid indicator of overweight and obesity for clinical, screening, and surveillance purposes and thus, has been utilized by the CDC as the criteria for which children and adolescents are categorized as overweight or obese (CDC, 2011; Himes, 2009). Additionally, BMI has demonstrated criterion validity in its association with several health outcomes (see Wabitch, 2000a; 200b).

However, there are practical limitations of the BMI when used to identify health status in children (Wheeler & Twist, 2010). Measurement error (and thus, a reduction in reliability), may result from practical issues related to the participant (e.g., clothing choice, hair style) and observer variation (Himes, 2009). Error in measurement may be reduced with training, as well as the use of more reliable and accurate electronic scales and stadiometers (Himes, 2009). Improved measurement methods may increase reliability and accuracy. For example, it has been suggested that multiple biometric readings be taken and averaged together for each measurement point (Himes, 2009). While the reliability of BMI depends on a multitude of factors, the reliability of BMI is generally high compared to other methods of body composition measurement (e.g., Skinfold thickness; Dietz & Bellizzi, 1999).

The clinical utility of BMI is high in a number of ways but it is not without limitations. BMI has high acceptability given its ubiquity in obesity-related research, namely currently being the generally accepted criteria for being overweight or obese. Feasibility of BMI is high compared to other health indicators (i.e., ADP, DXA, hydrodensitometry); however, practical constraints remain, such as required training to ensure reliability, increased time requirements, and labor force. BMI is also bidirectional and likely resistant to client desirability (if children are not provided with immediate feedback on their weight status). BMI as a numeric value has relatively low real-world meaning unless considered in percentile form compared to age-related norms. However, BMI as a tool to classify populations into levels of weight status (i.e., underweight, overweight, obese) is meaningful, despite the somewhat arbitrary nature of percentile cutoff scores. Similarly, criticisms of BMI often involve limitations in the BMI's ability to detect differences in body composition, such as density of different types of body tissue (Wheeler & Twist, 2010). For example, BMI has been found to be significantly correlated with percent body fat measured by air displacement plethysmography (ADP) in children, but with an "underwhelming" effect size (Pearson's $r = .45$; Wheeler & Twist, 2010). Thus, BMI is likely a helpful clinical screening tool, but intensive training to prevent low reliability and relatively low meaningfulness of index scores (compared to more face-valid outcomes such as body fat percentage) may limit the utility of BMI as an indicator of physical health in youth in certain contexts.

Skinfold Thickness. Skinfold thickness measurement is one of the most straightforward methods of determining youth body fat percentage. Special calipers

are used to measure the thickness of the subcutaneous fat layer at several sites of the body (i.e., triceps, biceps, subscapular). These values are then used to produce estimates of percent body fat using established equations based on gender, ethnicity, and stage of maturation (Slaughter et al., 1988; Steinberger et al., 2005). Gutin et al. (1996) investigated differences in body composition measures in 43 children (age 9 to 11) and found that skinfold thickness measurements had larger trial-to-trial differences of body fat compared to DXA and bioelectrical impedance (see below); however, the overall difference for all three methods was relatively low (e.g., the largest trial-to-trial difference for skinfold-thickness was only 2.8%). Furthermore, reliability as measured by intraclass correlation (ICC) indicated high reliability (ICC > .99). The construct validity of skinfold thickness has been supported by findings that indicate high correlations with body fat measurements from dual-energy X-ray absorptiometry ($r = .92$). Furthermore, skinfold thickness has been found to be indicative of cardiovascular risk, indicating criterion-related validity (Steinberger et al., 2005).

Regarding clinical utility, skinfold thickness calipers are generally accepted in assessing body fat; however, measurements require pinching of the skin in various parts of the body and may cause discomfort or anxiety in children compared to other measures such as BMI and bioelectrical impedance (see below). The feasibility of skinfold thickness is a strength in that calipers are generally low in cost (less than \$50 without software, \$200-400 with software) and portable. However, skinfold thickness measurements place a high demand on evaluators with regard to training (e.g., awareness of the proper location and angle on the body in order to ensure a proper

reading) and administration (e.g., high time consumption and mental demand; Amaral et al., 2010). Bidirectionality is good as skinfold thickness calipers allow one measure gains and losses in body fat, but reactivity to instrument use is possible given multiple administrations. The meaningfulness of degrees of change is easily understood given the physical measurement of fat on the body; however, calipers do not comprehensively measure body fat and equations must be used to translate measurements into body fat estimates. In general, skinfold thickness using calipers appears solid psychometrically, but has both strengths (e.g., low cost, portability) and weaknesses (e.g., evaluator demands) in terms of clinical utility.

Bioelectrical Impedance Analysis (BIA). BIA determines the electrical impedance (i.e., opposition to flow of electric current) through body tissue by utilizing the concept that fat-free mass contains nearly all of the body's conducting electrolytes (Tyrell et al., 2001). Biometric data such as height, sex, and age are used along with impedance information in a prediction equation in order to provide estimates of fat-free mass, total body water and body fat (Tyrell et al., 2001). There are multiple methods of attaining BIA data, including whole-body or segmental bioimpedance analysis. Whole-body bioimpedance analysis requires relatively expensive equipment (\$2,000 to \$5,000) but may be used to obtain raw resistance values to be used in published equations to determine body composition (Hannon, Ratliffe, & Williams, 2006). In contrast, segmental bioimpedance analysis uses relatively inexpensive equipment (<\$100) and "can only be used to obtain predicted estimates of body composition using unpublished equations, commonly referred to as 'proprietary' equations, developed and protected by bioimpedance manufacturers"

(Hannon, et al., 2006, p. 520). While some evidence indicates that whole body BIA yields higher estimates of body fat in girls (26.6% vs 23.2% in the same sample), it is difficult to interpret results since differences may be due to prediction equations utilized to estimate body fat instead of the actual measurements (Hannon et al., 2006; Slaughter et al., 1988). The potential differences between whole body BIA and segmental BIA measurements remain unclear; however, the expense of whole body BIA limits clinical utility in comparison with segmental BIA.

The two primary methods of segmental bioimpedance analysis are hand-to-hand (upper body) and foot-to-foot (lower body). The empirical support for hand-to-hand BIA is currently dubious. Hand-to-hand bioimpedance was found to be systematically lower than skinfold thickness measurements in a sample of Caucasian and African American youth; however, without knowing the true amount of body fat, it is difficult to compare the accuracy of these values (Hannon, et al., 2006). For example, hand-to-hand BIA may underestimate body fat while skinfold thickness may overestimate body fat. Gutin et al. (1996; see skinfold thickness section above) found high internal consistency and test-retest reliability for foot-to-foot BIA (ICC > .99; 2% or less change from trial 1 to trial 2). Body fat measured by foot-to-foot BIA has been found to have a very high correlation with a criterion measure of DXA, indicating convergent validity ($r = .98$; Tyrrell et al., 2001). While foot-to-foot BIA was found to be a better estimate of fat mass and body mass compared to other anthropometric indices (e.g., BMI), foot-to-foot BIA appeared to overestimate fat mass and body fat while underestimating fat-free mass (Tyrell et al., 2001). Consequently, foot-to-foot BIA may produce more false-positives in identifying

overweight and obese children. Thus, foot-to-foot BIA has limitations in specificity (i.e., the ability to eliminate false-positives), but has high sensitivity (i.e., the ability to correctly identify obesity).

Regarding the acceptability of BIA as a body composition measurement, Buchholz, Bartok, and Schoeller (2004) acknowledge BIA as one of the most commonly used body composition techniques in published studies and assert that BIA may be acceptable for determining body composition and monitoring change over time in groups. However, Buchholz et al. also cautioned against the utility of BIA in single measurements in individual patients due to the high likelihood of error. Additionally, BIA only requires children and adolescents to stand on scales, which is likely to be less intrusive and produce less discomfort compared to skinfold thickness measures or hydrodensitometry weighing. The feasibility of BIA is relatively high in that it is relatively inexpensive (e.g., less than \$100 per scale), portable, and requires limited training and time to administer. BIA allows for both gains and losses to be made; although, caution may be necessary when interpreting results at the individual level due to measurement error (Buchholz, 2004). Additionally, participant reactivity is possible if participants view their percent body fat while on the scale and thus, care should be taken to reduce this possible bias. The actual data measured by BIA is relatively meaningless in the real world and it is not until such data are translated through proprietary equations that the meaningful outcomes of total body water and body fat percentages are attained. The lack of meaning in BIA data is a primary limitation for an approach that, otherwise, demonstrates adequate reliability and validity and has high clinical utility in group settings.

Body composition measurement summary. Every body composition measure evaluated had strengths and/or weaknesses with regard to psychometric properties and clinical utility (see Table 2 and Figure 1). However, several themes emerged upon evaluation of youth body composition instruments. All of the instruments demonstrated some evidence to support their reliability and validity; although some had stronger evidence of their psychometric properties. In particular, ADP and DXA stood out for their stellar psychometric properties along with relatively good clinical utility with respect to acceptability, resistance to reactivity, and meaningfulness of results (Anderson, 2007; Fields et al., 2002). However, the cost of these instruments is likely to be a critical component in real-world decision making. Hydrodensitometry had one of the largest discrepancies between psychometric properties and clinical utility as the instrument is regarded as highly accurate; yet, frequently inappropriate for assessing children (Demerath et al., 2002; Reilly et al., 1995). Feasibility of use was a relative strength for BMI, skinfold thickness, and BIA in comparison to ADP, DXA, and hydrodensitometry. All measures demonstrated bidirectionality and most had low levels of reactivity. Most measures of body composition yielded body fat percentage as a health indicator through varying methods (BMI as a notable exception), which is helpful in determining meaningful changes while allowing for comparisons of the same outcome. While there is no standout “gold standard” measure of body composition in children and adolescents, ADP and DXA seem to maximize psychometric quality while BMI, BIA, and skinfold thickness maximize clinical utility.

Table 2

Validity, Reliability and Clinical Utility Criteria of Youth Body Composition Measures

Measure	Enlightenment		Clinical Utility				
	Reliability	Validity	Acceptability	Feasibility	Bidirectional	Repeatability	Meaningful data
Air Displacement Plethysmography (ADP)	Same day test-retest: $r = .98$; Coefficient of Variation = 5.3%	Very High	Easily administered; Potential discomfort of participants	Accommodates multiple populations; Expensive; Not portable	Yes	No reactivity	High: Body fat by subtraction
Dual-Energy X-ray Absorptiometry (DXA)	+/-2-4% between administrations	Very High	Often used as criterion method	Short duration; expensive equipment cost; lack of portability; required training in administration	Yes	No Reactivity; Ethical considerations in children	High: Body fat by subtraction
Hydrodensitometry	Coefficient of Variation = 6-10%	High	Limited in children (13-17% unwilling)	Specialized training and facilities required	Yes	No reactivity	Moderate to High: Body fat by equations based on hydration and body composition assumptions
Body Mass Index (BMI)	Generally high but prone to error without proper training	Moderate to High	Used as criteria by CDC for overweight/obesity; Commonly utilized	Training required for reliability; increased time and labor; portable	Yes	Reactivity is possible but preventable	Low to Moderate: BMI numerical value or percentile
Skinfold Thickness	ICC > .99; less than 3% between administrations	High	Commonly utilized; may cause discomfort or anxiety in participants	Low cost (< \$50 per caliper); portable; high training and administration demand	Yes	Reactivity is possible	High: Body fat percentage estimated by subcutaneous fat thickness
Bioelectrical Impedance Analysis (BIA)	ICC > .99; 2% or less between administrations	High (foot-to-foot BIA)	One of most commonly used; More accepted in groups vs individual contexts; Low participant burden	Low cost (< \$100 per scale); portable; low administration/training demand	Yes	Reactivity is possible but preventable	Moderate: Body fat percentage by ambiguous proprietary equations

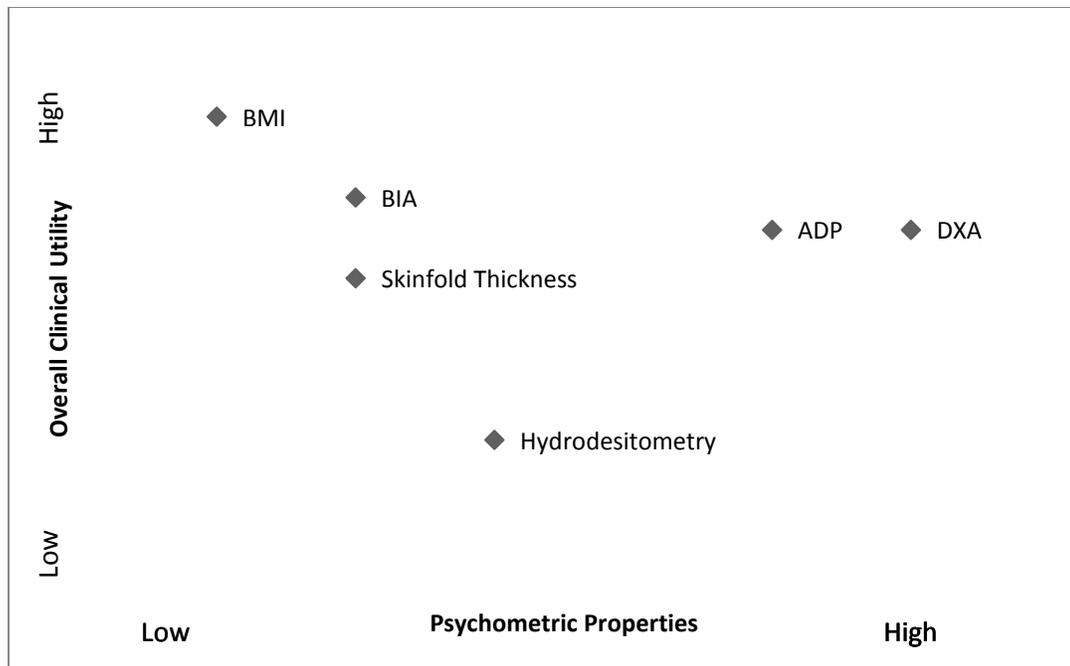


Figure 1. Psychometric properties and clinical utility of youth body composition measures. This figure represents subjective strengths and weaknesses of body composition measures in children and adolescents based on a balance of psychometric properties (i.e. “enlightenment”) and clinical utility (see Kazdin, 2005). Values are relative to other measures of body composition (e.g., measures closer to the low end may still be adequate in general but are relatively weak compared to other body composition measures).

Measuring youth physical activity. Measurement of youth physical activity is complex as the tempo of youth movement rapidly changes at intervals that are difficult to predict (Bailey et al, 1995). Subsequently, there are a number of methods of measuring physical activity in children and adolescents. The physical activity measures reviewed in the current study represent the most prominently studied measures based on searches for youth physical activity-related studies in several databases (e.g., ProQuest, SAGE, ScienceDirect, Wiley). Behavioral observation is presented first and, although multiple

methods of physical activity exist, the SOFIT in particular is discussed as it has been used as a criterion measure of physical activity in children. Next, DLW is discussed as a measure that is more commonly used in laboratory settings. Finally, pedometers and accelerometers are discussed as instruments that are more often used in field settings. Each instrument is reviewed systematically with regard to: enlightenment (reliability and validity) followed by clinical utility (acceptability, feasibility, resistance to reactivity, bidirectionality, and meaningfulness of data).

Behavioral observation. Behavioral observation has been suggested as the most appropriate measure of physical activity behavior in children, particularly in structured activity programs (Kohl et al., 2000). One primary method for assessing youth physical activity through behavioral observation is the System for Observing Fitness Instruction Time (SOFIT; McKenzie, Sallis & Nader, 1991). The SOFIT is a time sampling and interval recording system designed to assess level of physical activity, context of the lesson being observed, and teacher behavior during the lesson. The SOFIT measure uses prerecorded prompts to allow the observer to record the progression of movement and lesson context as it occurs (McKenzie, Sallis, & Nader, 1991). The level of physical activity is coded depending on body position of the student (1 = lying down, 2 = sitting, 3 = standing, 4 = walking 5 = very active i.e. expending more energy than ordinary walking). Moderate to vigorous physical activity (MVPA) is defined by a score of 4 (walking) or 5 (very active) on this measure. There are six options for the context of the lesson: Management, Knowledge, Fitness, Skill-building, Game, or Other (McKenzie, 2009; McKenzie, Sallis, & Nader, 1991). Management refers to time where physical activity or education is not intended. Included in this category would be activities such as

setting up equipment, taking attendance, dividing students into teams, transitioning teams or stations, and break time. Knowledge refers to time spent educating students in areas related to the lesson such as physical fitness, health, and rules of sports or games. Fitness refers to periods of time where activities are aimed at improving physical endurance, strength, or flexibility, including activities such as stretching, running laps, or aerobic dance. Skill-building refers to activities where the learning or development of a particular physical activity skill is the central focus, including activities such as dribbling a basketball, bumping a volleyball, or serving a tennis ball. Games are identified as activities where skills are applied in a competitive setting with minimal intervention of the instructor, including activities such as soccer games, tag, or kickball (McKenzie, 2009; McKenzie, Sallis, & Nader, 1991). Teacher behavior is also categorized into six options: Promotes fitness, Demonstrates fitness, Instructs generally, Manages, Observes, and Off-task. Initial evaluation of the SOFIT yielded high inter-rater agreement (approximately 88 to 92%).

Pope, Coleman, Gonzalez, Barron, and Heath (2002) investigated the reliability and validity of the SOFIT by comparing results to accelerometer-measured physical activity of 56 students during physical education lessons. Results indicated very high interobserver reliability of the SOFIT ($r > .90$), high internal consistency ($ICC = .98$), and increasing accelerometer-measured physical activity at increasing intensity of SOFIT levels (i.e. greater physical activity at level 3 versus level 4). The high interobserver reliability reported is an especially important finding for behavioral observation; where differences in observer characteristics can potentially lead to measurement error. Use of the SOFIT to estimate physical activity has been validated empirically in children and

adolescents, with some evidence indicating that a six-point scale may be more sensitive to measuring variations in light to moderate physical activity (Heath, Coleman, Lensegrav, & Fallon, 2006; McKenzie et al., 1991; Pope et al., 2002). Thus, the SOFIT appears to demonstrate high reliability and construct validity (i.e., that the SOFIT is measuring physical activity as intended) in a physical education setting.

Regarding clinical acceptability, the SOFIT has been utilized as a criterion measure for other methods of physical activity measurement in children, particularly in structured activities times. In 2005, Scruggs, Beveridge and Clocksin compared behavioral observation to tri-axial accelerometry and heart rate telemetry in elementary physical education. Authors utilized heart rate monitors, accelerometers, and the SOFIT as their behavioral observation measure in order to calculate and compare estimated MVPA from each measure. Scruggs et al. (2005) utilized the SOFIT as their criterion measure based on the assumption that behavioral observation is the most accepted method of recording physical activity data in children. Results indicated that in comparison to SOFIT-measured MVPA, heart rate monitor and accelerometer-measured MVPA “would not be considered clinically acceptable for accurately measuring time engaged in physical activity” (p.214). Heart rate monitor results in particular had only moderate correlations with behavioral observation data ($r = .42$ to $.49$). While such results support the clinical acceptability of behavioral observation data; conclusions regarding accelerometry and heart rate monitoring as unacceptable may be questionable. For example, heart rate continues to be elevated after physical activity has been completed, which could have inflated MVPA estimates, leading to reduced correlations with observed MVPA. Correlations between accelerometer-estimated MVPA and

behaviorally observed MVPA were higher ($r = .77$ to $.78$), yet authors indicate that accelerometry systematically overestimated MVPA and thus is not an acceptable alternative for behavioral observation. However, authors acknowledged that the cutoff they used for movement to be considered MVPA may have been set too low for accelerometry, which could have inflated their accelerometer-estimated MVPA. Furthermore, accelerometers were set to record in 60 second intervals which may result in a lower association with behavioral observation than a shorter epoch time (e.g., 1, 5 or 20 seconds).

Regarding clinical utility and, in particular, feasibility, the SOFIT requires proper observer training including videotaped training sessions, multiple field trials, and multiple observers to establish inter-rater reliability (McKenzie, 2009; Pope et al., 2002). Thus, the utility of behavioral observation is reduced in large sample due to large time and personnel requirements, as well as subsequent cost. While MVPA is, to an extent, a bidirectional measure of fitness (e.g., minutes or percentage of time engaged in MVPA may increase or decrease); ceiling effects are possible (i.e., scores so high that improvement is difficult to measure); which is exacerbated when “moderate” cutoffs for MVPA are utilized (i.e., when moderate and vigorous physical activity are equated). For example, if a student briskly walked for the duration of their 20 minute activity period, improvement would not be measured by the SOFIT if they were to run wind sprints over the same period of time (and may in fact show reduced performance if they slowly walked between sprints). Reactivity to the SOFIT is a possibility and counter-measures should be taken. Observers are taught during training to observe globally (in lieu of

looking directly at the target student) and to avoid specific feedback and discussions about observations with students and teachers (McKenzie, 2009).

The meaningfulness of SOFIT-estimated MVPA is relatively high since youth are actually seen performing the physical activity (i.e., highly face valid). Results are even more meaningful in that the SOFIT allows for other elements such as lesson context (e.g., skill-building, fitness, game) to be considered along with physical activity. Some precision and explanatory power is lost when data is transferred into MVPA due to loss of degrees of change (i.e., a continuous five or six-point scale dichotomized into MVPA or no MVPA). However, MVPA is a ubiquitous outcome measure in physical activity research and translation of SOFIT data into MVPA does ease comparisons of youth physical activity in the literature. Another drawback in meaningfulness is that momentary time sampling estimates behavior by sampling behavior at the end of a specified period of time. Combined with the erratic nature of physical activity in children and adolescents, proper time sampling is essential in order to provide a meaningful representation of physical activity during the observation period. In 2005, McNamee and van der Mars (2005) investigated the accuracy of several lengths of sampling time in assessing youth physical activity. As expected, error increased with longer time samples, and 20 second intervals appeared to serve as the ideal length of time between samples. Furthermore, authors found 90 second intervals to be the longest length of time that allowed them to accurately measure youth MVPA. Another generally important consideration for the clinical utility of behavioral observation is the context of the measurement. As discussed, there is evidence to support its use in structured physical activity time; however, behavioral observation is likely to be less reasonable when the aim is to measure physical

activity in free-living conditions (i.e., energy expended in everyday living). Increased time, cost, likelihood of reactivity, and ethical considerations would likely limit practical use in attempting to capture youth physical activity over extended time periods and multiple locations. In general, behavioral observation (i.e., SOFIT) is perhaps the most accepted measure of physical activity in structured, time-limited contexts (e.g., physical education, after school activities, sport activities). However, the clinical utility of behavioral observation is primarily limited by burden placed on evaluators and contexts for which the instrument's application is reasonable (e.g., not in free-living situations).

Doubly Labeled Water (DLW). DLW is a noninvasive method of measuring daily energy expenditure in children and adolescents, which can be combined with resting energy expenditure to determine energy expended through physical activity. The DLW procedure begins by having the participant ingest a known volume of water containing two isotopes of water, $^2\text{H}_2\text{O}$ (deuterium-labeled water) and H_2^{18}O (oxygen-18-labeled water). The difference in the rate of loss of the two isotopes is indicative of physical activity as oxygen-18-labeled water is partially eliminated via carbon dioxide production while deuterium-labeled water is lost from the body at the natural elimination rate (Goran, 1994; Schoeller, 1988). Research regarding reliability of doubly labeled water indicated 8.5% to 12% variance between controlled administrations (Goran, Poehlman, & Danforth, 1994). There is also evidence to support the construct validity of DLW (within 2 and 8% of energy expenditure determined from monitored dietary energy intake and body composition) in measuring energy expenditure (Schoeller, 1988).

With respect to the clinical utility, DLW appears to be a relevant, easily administered procedure; however, the procedure requires extensive expertise in analysis.

Additionally, DLW is applicable in multiple contexts including free-living situations. However, the feasibility of DLW is limited by high cost of materials (e.g, isotopes), equipment, and expert analysis (Trost, 2001; 2007). Data measured by DLW as energy expenditure is bidirectional and likely to retain validity over time (Schoeller, 1988). The real-world meaningfulness of doubly labeled water data is dependent on the way in which the data is presented. Initial measures seem low given the complexity of the procedure and data outcome (i.e. difference in isotope rate; carbon dioxide production); however, once translated into daily energy expenditure (ideally, kilocalories or Calories) the real-world meaning is much improved. Additionally, the total energy expended after an extended period of time lacks other elements that may be important in studying physical activity such as context, type, or intensity of movement. In other words, there is no information on how the energy was expended (e.g., a long walk may expend the same amount of calories as a shorter run). While the psychometric properties of DLW are impressive, the high cost of materials, expertise required in analysis and inability to provide data on the context of physical activity remain significant drawbacks and consequently may limit clinical utility.

Pedometer. Pedometers are small, portable devices that are placed on the hip in order to count the number or steps taken while walking or running (Tudor-Locke, Williams, Reis, & Pluto, 2002). The majority of modern electronic pedometers work through opening and closing an electrical circuit via a spring-suspended lever arm that deflects upon the up-and-down movement of the hips during ambulatory activities (Tudor-Locke et. al., 2002). Evaluations of the reliability of pedometers at varying degrees of intensity yielded intra-class correlations (ICC) from $r = .52$ to $.92$ in a sample

of 78 adolescents (Jago et al., 2006). A study evaluating the accuracy of pedometer step counts in children using treadmills indicated a high level of accuracy and step count agreement among pedometer models at or above common walking speeds (i.e., above 2.5 miles per hour) but low agreement during slow-paced walking (i.e., less than approximately 2 miles per hour; Beets, Patton, & Edwards, 2005). In 2002, Tudor-Locke et al. performed an evaluation of the convergent validity of pedometers and found that pedometers may be an acceptable, simple, and cost-effective alternative to more burdensome and/or expensive methods when the intent is to measure ambulatory movement. Pedometers were found to correlate highly with accelerometers ($r = .86$), with particularly high correlations among uniaxial pedometers. High correlations were also found with observed activity time ($r = .82$); however, there was consistent evidence that pedometers have reduced accuracy during slow walking.

The acceptability of pedometers is relatively high due to low participant burden (i.e., wearing a small device at the hip) and ease of administration. Feasibility is also a relative strength of pedometers as a result of the low cost, portability, and user-friendliness of the instrument. Furthermore, pedometers may be utilized in a multitude of contexts (e.g., physical education periods, free-living conditions, large-scale studies). Step counts as measured by pedometer are bidirectional to an extent (i.e., the participant either took more or fewer steps) but may not necessarily represent gains and losses due to lack of movement intensity (Troost, 2007; Tudor-Locke et al., 2002). In other words, the step count, provided by pedometers may not be representative of energy expenditure since all steps are not necessarily equal. For example, a pedometer would not be able to distinguish between one hundred steps casually walking from class to class from one

hundred steps while running at maximum speed during football practice. Thus, data garnered from pedometers may be meaningful in real-world contexts, but may not be an “accurate” measure of physical activity. Interestingly, the face-validity of step counts may impact validity over time. Since most children and adolescents can understand that the device is simply counting their steps, the instrument is vulnerable to reactivity (Troost, 2007), which is likely to be exacerbated if the pedometer model displays step counts (providing immediate feedback on progress). The pedometer is a psychometrically adequate and cost-effective measure of static ambulatory movement (i.e., continuous walking or running), but not when changes in movement intensity or non-ambulatory movements are intended to be addressed in the study (Tudor-Locke et al., 2002).

Accelerometry. Accelerometers are electronic devices attached to the body, typically the hip or lower back, in order to provide quantitative information regarding body accelerations at specified time intervals, called epochs. Accelerometers have similarities with pedometers in that they are both devices placed on the body to measure activity; however, accelerometers provide a more dynamic picture of movement as they are able to measure the intensity of movements at a specified time period (e.g., every minute) instead of a total number of step counts over the entire measurement period. Additionally, triaxial accelerometers (essentially three uniaxial accelerometers combined in one device) are able to provide information on movements along three planes (i.e., up-and-down, side-to-side, forward-and-backward) instead of merely up-and-down motion. Accelerometer data is recorded in “counts” over a given period of time which are then entered into equations to yield more interpretable measures of physical activity such as MVPA or energy expenditure (Troiano, 2006).

The use of accelerometers as a measure of energy expenditure has been investigated in a number of studies with the majority of research showing strong correlations between accelerometer measurements and energy expenditure or exercise intensity as measured by heart rate telemetry, doubly labeled water, and behavioral observation (Freedson, Pober, & Janz, 2005; Janz, 1994; Trost, McIver, & Pate, 2005). The use of triaxial accelerometers to assess energy expenditure in children has also been validated in simulated free-living conditions such as sitting, writing, laying down, cycling, stepping, jogging, and performing tasks related to basketball, soccer, and tennis (Sun, Shmidt, & Teo-Koh, 2008). One study completed by Scruggs et al., (2005) yielded contrary evidence as accelerometry systematically overestimated MVPA in comparison to behavioral observation in a sample of 346 first- and second-grade students. Authors concluded that accelerometry is not an acceptable alternative to behavioral observation in estimating physical activity time in children. However, accelerometry and direct observation were still highly correlated ($r = .77$ to $.79$) and several methodological flaws may have affected the accuracy of accelerometer-determined MVPA. First, the study used 60-second time intervals for their accelerometer recordings while current accelerometers have the ability to be record movement as frequently as one-second, which would be likely to increase accuracy. Furthermore, authors acknowledged that the cut-point values that were used to translate raw accelerometer counts into MVPA were too low, leading to overestimated MVPA. Despite methodological limitations, results suggest that behavioral observation may be the preferred method of physical activity measurement compared to accelerometers during physical education lessons. However, youth physical activity does not take place exclusively during physical education and it is

likely impractical to use behavioral observation methods to measure physical activity in free-living situations. One other noted limitation of accelerometers is that they are not able to capture increased energy required to move up stairs or on an incline or lifting and carrying objects; however, the contribution of such activities is likely to be small in overall energy expenditure in children (Freedson et al., 2005; Trost, 2007; Welk et al., 2000).

The ability of accelerometers to record and store data on the device allows for measurements to be taken for days at a time, allowing for the estimation of typical daily or weekly MVPA time. In order to appropriately estimate energy expenditure or activity intensity using accelerometers, great care must be taking in selecting the product, placing the device on and preparing participants, setting appropriate time intervals (epochs), setting appropriate cut points for physical activity, and allowing a sufficient number of monitoring days (Trost et al., 2005). The most popular models of accelerometers used in research have been the ActiGraph, Actical, and RT3 triaxial accelerometers. The ActiGraph accelerometer is small, lightweight, and the most widely used monitor (ActiGraph LLC, Fort Walton Beach, FL; Trost, 2007). Advantages of the ActiGraph include direct USB connection, 1Mb unit memory, self-calibrating digital accelerometer, and documented validity in children and adolescents (Freedson et al., 2005; Trost et al., 2005). The Actical accelerometer is unique in that it is the smallest available accelerometer, contains an omnidirectional sensor, and is water resistant (Mini Mitter Respironics, Bend, OR; Trost, 2007). An investigation of the validity of Actical counts reported high correlation with VO_2 ($r = .89$) and high sensitivity (97.2%) and specificity (91.7%) for estimating energy expenditure (Pfeiffer, McIver, Dowda, Almeida, & Pate,

2006). The RT3 Accelerometer is a newer, slimmer model of the Tritrac R3D, which was the first commercially available accelerometer with the ability to measure movements along three axes (Stayhealthy Inc., Monrovia, CA). Recent research in youth simulating free-living conditions utilized the RT3 and suggests that it is able to provide acceptable estimates of physical activity through cross-validation with alternative measures of energy expenditure ($r = .77$ to $.98$; Sun et al., 2008). Review of comparison studies of different accelerometer models garnered inconsistent results, with moderate ($r = .52$) to high ($r = .92$) correlations between most models, and no particular product emerging as being superior (see Trost et al., 2005). Overall, the lack of consistency of results in comparison studies leaves researchers to consider other important factors when making decisions about “make and model” (e.g., cost, device size, or use in studies of similar design).

The placement of the accelerometer has been investigated in multiple studies with general findings suggesting that the device is best placed on the hip or lower back (Trost et al., 2005). A study investigating physical activity in children over a span of four days found no differences in counts per minute or MVPA time when comparing hip and back placement (Nilsson, Ekelund, Yngvie, & Sjostrom, 2002). The choice of epoch time and cut point for MVPA has been found to have a significant relationship with “sedentary time”, with shorter epoch times and less strict cut points for moderate activity resulting in lower estimated sedentary time (Ojiambo et al., 2011). Recent research indicates that the reliability of accelerometers in assessing free-living physical activity ranges from very good to excellent (intra-class correlation $r = .70$ to $.90$), but vary depending on algorithm used to calculate physical activity (Sirard, Forsyth, Oakes, & Schmitz, 2011).

Furthermore, several equations have been developed to predict energy expenditure using accelerometer counts but no equation has emerged as universally appropriate and the choice of equation can greatly overestimate or underestimate energy expenditure (Mota et al., 2007; Nilsson et al., 2008; Trost, 2007). It has been hypothesized that difficulty in establishing appropriate cut points and estimating equations in practical application may be due to equations being created using non-representative samples in laboratory or simulated free-living conditions instead of actual living conditions (Mota et al., 2007; Ojiambo et al., 2011).

Regarding clinical utility, Accelerometry can be considered to be acceptable in assessing physical activity in children as it expands upon the limitations of pedometer use with regard to movement intensity. However, great care and consideration is required during administration and interpretation procedures to produce relevant results (e.g., “make and model,” cut points for MVPA or energy expenditure). Feasibility is a relative strength during administration given the small size, light weight, durability, and moderate cost (approximately \$200-\$300 per unit); however, accessing results is an obstacle due to the expertise, time, and cost of carefully analyzing data (Trost, 2007; Welk, Corbin, & Dale, 2000). The counts recorded by accelerometers are bidirectional as is the translated result in energy expenditure. Accelerometry is improved upon its pedometers in resistance to reactivity in that movements recordings are more ambiguous compared to step counts. However, the need to transfer accelerometer counts into a more interpretable unit is a primary limitation of accelerometry in comparison to measures of physical activity that are more meaningful and interpretable in their outcomes (e.g., behavioral observation). Accelerometers overall seem to have generally strong psychometric

properties and clinical utility when care is taken in the decision-making process (e.g., epoch length, placement, “make and model”); however, difficulties in interpreting results due to ambiguous recorded values impair clinical utility.

Physical activity measurement conclusions. Like measures of body composition, measures of physical activity had strengths and weaknesses with regard to psychometric properties and clinical utility (see Table 3). Interestingly, measures with better psychometric properties generally appeared to have lower clinical utility (see Figure 2), a pattern that was not as prominent in body composition measures. One possible explanation for this observation has to do with the erratic and mercurial nature of physical activity in children (Bailey et al., 1995). Perhaps difficulty measuring physical activity leads to much greater effort and expense in order to attain high validity and reliability. For example, behavioral observation and DLW exhibited the strongest psychometric properties but behavioral observation requires skilled labor, extensive training, and thus considerable time and cost (McKenzie et al., 1991; Pope et al., 2002; Scruggs et al., 2005) while DLW requires expensive materials and expertise to execute the complex task of interpreting results (Goran et al., 1994; Trost 2001; 2007). Pedometers seem to be on the opposite end of the spectrum and are primarily strong in aspects of clinical utility such as acceptability and feasibility (e.g., ease of use). The clinical utility of pedometers may in fact be linked with their psychometric properties since the limitations in face-validity of step counts may cause reactivity and compromise validity. Accelerometry seems to have evidence of better psychometric support than pedometers; however, clinical utility is reduced due to lower feasibility (e.g., increased cost, and required expertise in interpreting data) and meaningfulness of the data recorded. No single measure emerged

as a “gold standard” measure of youth physical activity; however, DLW and behavioral observation appeared to maximize psychometric quality while pedometers and accelerometers had greater strengths in overall clinical utility (see Figure 2).

Table 3
Validity, Reliability, and Clinical Utility of Physical Activity Measurements

Measure	Enlightenment		Clinical Utility				
	Reliability	Validity	Acceptability	Feasibility	Bidirectionality	Validity over time	Meaningful data
Observation (SOFIT)	Inter-rater reliability $r > .90$	High	Common criterion measure; Able to evaluate context and intensity; Less practical in free-living conditions	High labor cost; Low cost of materials; Extensive training; Less practical in large samples	Improved in five or six-point scale compared to all-or-nothing MVPA	Reactivity is possible; Prevention procedures during training	High as movement intensity; Moderate as MVPA but useful in research
Doubly Labeled Water	Coefficient of Variation = 8.5%	High	Easily administered; Utility in free-living conditions; Expertise required	Expensive materials, equipment, and labor; Less practical in large samples	Yes	Low reactivity	Low in raw data; High in kilocalories per day; Lacks context and intensity of activity
Pedometer	ICC = .52 to .92	Good	Easily obtained and administered; Utility in free-living conditions	Low cost; Lacks context or intensity; Portable; Relies on adherence; Utility in large samples	Yes	Vulnerable to reactivity	Moderate to High: Step counts; Lacks context and intensity
Accelerometry	ICC = .70 to .90	Good	Easily administered; Utility in free-living conditions; Expertise required to interpret results	Moderate cost; Lacks context; Portable; Relies on adherence; Utility in large samples	Yes	Limited vulnerability to reactivity	Low: ambiguous data counts must be translated using equations to estimate more relevant outcomes

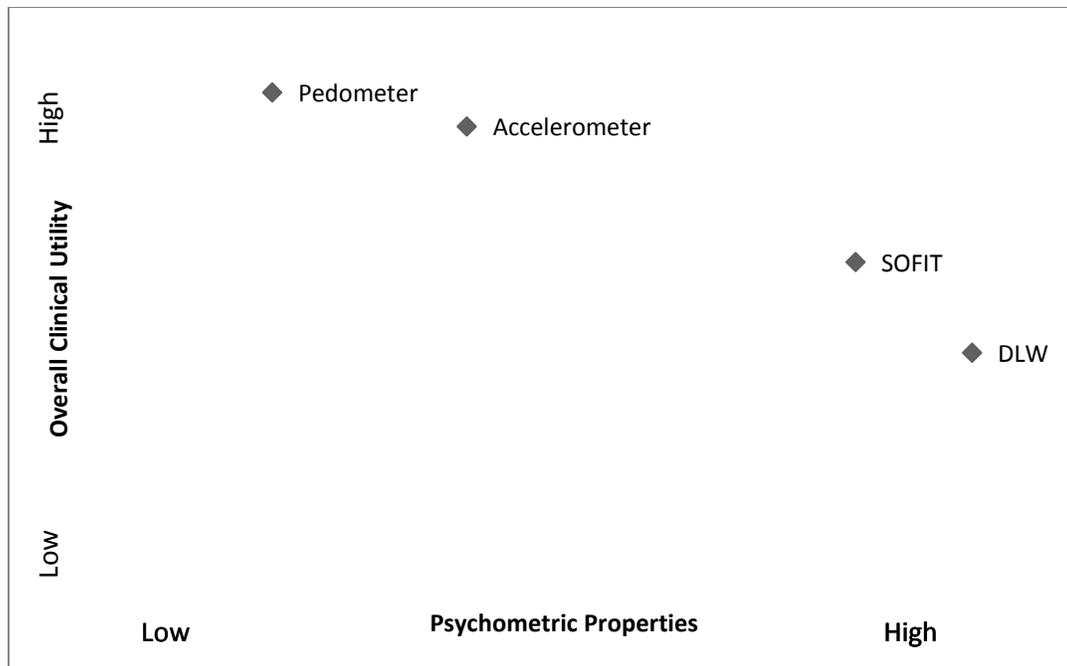


Figure 2. Psychometric properties and overall clinical utility of physical activity measures in children and adolescents. This figure represents subjective strengths and weakness based on enlightenment (reliability and validity) and clinical utility (see Kazdin, 2005). Values are relative to the instruments displayed (e.g., an instrument may have adequate psychometric properties or clinical utility but be relatively weak in comparison to other physical activity measures).

Clinical utility of body composition and physical activity. In general, measures of physical activity seemed to have adequate psychometric properties but were limited psychometrically when compared to measures of body composition (see Tables 1 and 2). In general, clinical utility was higher for physical activity measures compared to body composition measures. Most measures were relatively acceptable when the measurement was used in the appropriate context (e.g., SOFIT in lesson contexts, DLW and Accelerometry in free-living conditions). Feasibility varied greatly from one instrument to another, particularly when cost (e.g., ADP, DXA and DLW expensive, skinfold

thickness calipers and pedometers inexpensive) and evaluator burden (e.g., high burden in observation, low burden in pedometers) are considered. Strength in bidirectionality was a consistent finding across body composition and physical activity measures; however, some physical activity measures (e.g., SOFIT, Accelerometer) may represent data as MVPA, which dichotomizes data in a way that may obscure more subtle changes in activity level. Reactivity to measurements was a relevant issue for more of the physical activity measures, though simple changes in procedure may counter this bias (e.g., observing globally using the SOFIT; utilization of pedometers without visible step counts). Physical activity measures varied in the meaningfulness of outcome data (i.e., MVPA, energy expenditure, movement counts), especially compared to body composition where a general consensus was apparent (e.g., percent body fat). For both body composition and physical activity, no singular measurement stood out as a “gold standard” instrument by maximizing both psychometric properties and clinical utility. In general, body composition measures appeared to be more psychometrically robust (i.e., “enlightened”) while physical activity measures seemed to have greater clinical utility (Kazdin, 2005).

The “gold standard” problem and multiple instrument utilization. Kazdin (2005) discussed the issue lack of a “gold standard” as a common theme that must be considered in child and adolescent assessment. Kazdin suggested that there is no one measure that captures a given clinical problem and suggests that multiple measures are needed to evaluate different facets of the problem. The reviews of body composition and physical activity measurements in the current study align with Kazdin’s assertions and indicate that currently, there is no single measure of body composition or physical activity that

maximizes psychometric quality and clinical utility to a point that it is sufficient alone as a health indicator. Instead, evaluators should consider the relative strengths and weaknesses regarding psychometric properties and clinical utility and choose multiple instruments that are appropriate for their given purpose. For example, a study that aims to evaluate a physical education program would likely benefit from utilizing the SOFIT as a direct observation measure to assess physical activity during daily lessons. However, physical activity does not only take place during lessons and measurement of free-living activity would be helpful in determining whether or not the physical education program improves physical activity at home. Unfortunately, the acceptability and feasibility of observation instruments for measuring free-living physical activity is limited. Thus, future research may benefit from more widespread use of accelerometry along with the SOFIT in order to address the SOFIT's limitations and provide more accurate assessment of the overall effectiveness of the physical education programs. Therefore, instead of seeking a single youth health indicator, researchers should consider the way in which such measures may complement one another.

The Relationship between Body Composition, Physical Activity, and Physical Fitness

While the health benefits of increased physical activity and negative health effects of obesity seem obvious, research on the relationship between body composition and level of youth physical activity has yielded inconsistent results. Rennie et al. (2005) investigated the relationship between physical activity and body composition in 100 children at varying risk of obesity (e.g., none, one or both parents with obesity) using DLW-calculated energy expenditure and heart-rate measured physical activity. Results

indicated that the high-risk children had higher BMI and fat mass index but there was no difference in physical activity between risk groups. Physical activity and energy expenditure were positively correlated with lean mass index and negatively correlated with fat mass index (but not BMI) after adjusting for gender and fat-free mass.

Additionally, boys that spent more than 36.6% of their time in light-intensity activities had higher fat mass index than less sedentary boys, but this finding was not found in girls and no associations between vigorous activity and body composition were found.

Abbott and Davies (2004) investigated the relation of physical activity intensity to body composition in 47 Australian children using DLW, accelerometer-estimated MVPA, foot-to-foot BIA and BMI. Children wore triaxial accelerometers for a four-day period (split evenly between weekday and weekend), in order to assess habitual physical activity. Results indicated that body fat and BMI were negatively correlated with physical activity level measured by DLW ($r = -.45$ to $-.43$), and time spent in vigorous or higher (but not moderate) activity was also negatively correlated with BIA ($r = -.44$). Such results indicate a moderate association between habitual vigorous physical activity and body composition. While these results suggest a relationship between body composition and physical activity, findings across studies are inconsistent and interpretations are further complicated by variance in methodology.

In 2005, Gutin et al. investigated the relationship between intensity of physical activity to fitness and body fat in a sample of 421 high school adolescents. Investigators used age, race, gender, and interactions as covariates then investigated the influence of accelerometer-measured physical activity (categorized into moderate, vigorous, and MVPA) in order to predict DXA-measured body fat and cardiovascular fitness measured

by a multistage treadmill test. Results indicated that only vigorous physical activity significantly predicted body fat percentage over and above demographic variables (total $R^2 = .38$) while moderate, vigorous, and MVPA all contributed significantly to cardiovascular fitness when accounting for demographic variables (moderate $R^2 = .37$; vigorous $R^2 = .42$; MVPA $R^2 = .38$). While this finding is important in demonstrating the impact of physical activity when accounting for other variables, it is perplexing that the two outcome variables of body fat and fitness were analyzed separately and were never used in models to predict one another. Thus, questions remain regarding the relative contribution of body composition and physical activity in accounting for youth cardiovascular fitness.

A study conducted in Portugal by Aires et al. (2010) investigated the relationship among BMI, accelerometer-measured physical activity, and cardiorespiratory fitness as measured by a 20 meter multi-stage shuttle run (Progressive Aerobic Cardiovascular Endurance Run [PACER]) in a sample of 111 students in age 11 to 18. Investigators used a logistic regression to predict overweight/obesity status (adjusted for age and gender) by physical activity intensity and cardiorespiratory fitness. Results indicate that only cardiorespiratory fitness significantly predicted weight status (OR = .968, $p = .037$). While this result at first seems to suggest that physical activity is not associated with weight status when accounting for physical fitness, further examination may be necessary to clarify this relationship. Despite the lack of statistical significance, effect size presented as odds ratios indicate that those who engaged in vigorous physical activity were 37% less likely to be overweight or obese and those who engaged in very vigorous physical activity were 52% less likely to be overweight or obese as compared to those

with increased cardiorespiratory fitness being only 3% less likely to be obese. The lack of statistical significance may be limited by increased variation in accelerometer counts compared to PACER laps and may reach significance given increased sample size. Other limitations include the lack of indicators of variance accounted for (i.e., R^2 values), dichotomizing continuous variables (i.e., BMI into two weight status categories), and a lack of consideration of ethnic categories.

Lohman et al. (2008) investigated the influence of body composition and physical activity on physical fitness using a sample of 1148 eighth grade girls. Investigators utilized a regression model using body composition (weight; fat-free mass and fat mass calculated by skinfold thickness calipers), racial group (Non-Hispanic Black and all others) with body composition interaction effects, and accelerometer-measured daily MVPA to predict cardiorespiratory fitness measured by workload during a pedaling task. Results indicated that daily MVPA accounted for approximately 4% of the variance of fitness over and above body composition, race/interactions, and the other treatment variables (see Table 4). Overall, the model accounted for approximately 22% of the variance in cardiovascular fitness. Furthermore, results suggested that Black adolescent girls have lower fitness levels as well as physical activity levels compared to White adolescent girls and that the racial group x fat-free mass interaction term accounted for 4% of the variance of fitness before considering physical activity (Lohman et al., 2008; see Table 4).

While this finding is important in determining the relative contribution of body composition and physical activity in predicting physical fitness, several limitations are also present. First, generalizability of results are hindered by the restricted age and gender

of participants (i.e., only eighth grade girls). There is empirical evidence indicating that gender and age are significant influences of physical activity, body composition, and cardiovascular fitness and that gender differences generally seem to increase throughout adolescence (Keller, 2008; Troiano et al., 2008). Future studies evaluating the relative influence of variables on physical fitness in children and adolescents would provide more relevant information by utilizing a wider age range and both genders. Furthermore, Lohman et al. (2008) indicated that in their sample the Non-Hispanic Black group was the only ethnic group that was different from all others in fitness, leading to participants being dichotomized into two categories (Non-Hispanic Black or all other ethnicities) in their regression model. This limits the ability to determine the relative influence of other ethnicities such as Hispanic or Non-Hispanic White. While Non-Hispanic Black girls seem to be at particular risk for obesity, this ethnic disparity is not consistent across ethnic-gender subgroups (e.g., Hispanic boys are at elevated risk compared to Hispanic girls; Ogden et al., 2010).

Table 4
Hierarchical Regression Results Predicting Fitness in 8th Grade Girls (Lohman, 2008)

Step	Variable added to model	Overall R ²	Change R ²
	Basic Model (Location, Intervention Group,		
1	Racial Group)	0.05	0.05
2	Weight	0.1	0.05
3	Fat-free Mass, Fat Mass	0.14	0.04
4	Racial group x Fat-free Mass	0.18	0.04
5	Daily MVPA	0.22	0.04

The current study aims to utilize similar methodology to that performed by Gutin et al. (2005), Aires et al. (2010), and Lohman et al (2008) but with several key

improvements. First, increased sample size, age range, and ethnic diversity will increase power as well as external validity. Second, relationships will be analyzed with fitness (measured by the PACER 20-meter shuttle run) as the primary outcome variable. The rationale for this is to determine the impact of physical activity above and beyond other participant characteristics, including demographic variables and body composition. Other improvements of the current study include the utilization of a shorter epoch time in accelerometer readings (3 seconds compared to 1 minute), consideration of percent body fat as well as BMI, and consideration of gender/race interactions.

The role of demographic variables. Findings of the relationship among body composition, physical activity, and physical fitness are further complicated by important demographic factors such as gender, age, and race/ethnicity. In 2008, Troiano et al. evaluated data from 6329 children and adolescents who provided at least one day of accelerometer data as part of the 2003-2004 NAHNES (CDC). Results indicated that overall, males were more physically active than females and that there is a substantial decline in physical activity from age 6 to 19. One follow-up study confirmed such results and indicated that the steep decline in physical activity through adolescence occurs regardless of weight status (Chung et al., 2012). Investigators indicated that adolescent girls spent less than 5 minutes daily in vigorous physical activity while boys spent less than 8 minutes daily in vigorous physical activity, regardless of weight status.

While the impact of age and gender on physical activity and fitness seems to be relatively consistent in the literature, the influence of race/ethnic groups is less clear. There has been empirical evidence indicating that there are no differences in habitual physical activity between Black and Non-Hispanic White groups as reported by parents;

however, Black students reported less physical activity during physical education periods, more television viewing time, and less sports team participation (Lindquist, Reynolds, & Goran, 1999). Contrarily, studies using accelerometers to assess physical activity in a sample of over 1500 sixth grade females has indicated that Non-Hispanic White adolescent girls have more MVPA compared to Black and Hispanic girls (Lohman et al., 2008; Pate et al., 2006). However, another study assessed over 2000 European children aged 9 to 10 using accelerometers and found that Non-Hispanic White children are less physically active compared to other race/ethnic groups (Owens et al., 2009). Furthermore, another study using NHANES data to evaluate the impact of race/ethnicity on physical activity ($N = 3106$; age 6-19) found that Non-Hispanic Black youth recorded more time in MVPA (60.2 minutes/day) compared to Non-Hispanic White (52.3 minutes/day) and Mexican American youth (57.7 minutes/day; Belcher et al., 2010). The current study aims to clarify such discrepancies using accelerometer-measured physical activity in a diverse sample of children and adolescents.

Differences between race/ethnic groups in fitness levels have also been inconsistent among empirical studies. For example, Beets and Pitetti (2004) investigated the cardiovascular fitness of youth using a one-mile run/walk measure in an extremely large and diverse sample of 767,809 youth age 10 to 15. Results indicate that Non-Hispanic Black and Hispanic teenagers were consistently behind Non-Hispanic White teens in both males and females, with the effect becoming more pronounced with an increase in age. One setback of this study was that the data was presented only with respect to Non-Hispanic White fitness levels (i.e. no statistical comparison was made between ethnic minority groups). As measured by maximal VO_2 during a progressive

treadmill test, Lindquist et al. (1999) did not find any differences in physical fitness regarding race/ethnic categories in children aged 6 to 13 after controlling for fat mass and fat-free mass. Such contrary findings in make it difficult to draw definite conclusions about the role of race/ethnicity in fitness and indicate a need for future studies to further clarify race/ethnicity differences in youth fitness.

As obesity continues to be present at alarming rates in U.S. youth, there is a need to fill gaps of knowledge in physical activity and fitness research in order to better understand and improve associated health outcomes. Due to difficulty in measurement, differences in research methodology, as well as conflicting conclusions between studies, lingering questions remain regarding the relationship between body composition, physical activity, and cardiovascular fitness. The relationship among these concepts is further complicated by influential moderating variables such as gender, and age. Furthermore, the role of race/ethnicity continues to remain unclear. The present study is seeking to assist in clarifying the relationship between body composition, physical activity, and cardiovascular fitness while considering important demographic variables such as gender, age, and ethnicity. Follow-up analyses aim to clarify differences between ethnic groups in physical activity.

CHAPTER II: METHOD

Participants

The current sample included 58 students aged 6-17 enrolled in out-of-school programs in Miami-Dade County as part of larger study evaluating physical activity in out-of-school programs. Participants were randomly assigned to receive accelerometers by identification number. In the final sample, 81% of the sample identified as Hispanic, and 19% as Non-Hispanic Black. Similar to previous findings, more than half (59%) of the overall sample could be categorized as overweight, with Hispanic boys and Non-Hispanic Black girls at particular risk for being overweight (see Table 5). Overall the average percent body fat as measured by BIA was 26%, and more than half of all participants had an income of \$25,000 per year or less (see Table 5).

Table 5
Descriptive Statistics of Sample

	Hispanic (<i>n</i> = 47)	Non-Hispanic Black (<i>n</i> = 11)	Total (<i>N</i> =58)
Mean Age	9.9	10.0	9.9
BMI above 85th Percentile (overweight)	60%	55%	59%
Boys (<i>n</i> = 27)	66%	50%	63%
Girls (<i>n</i> = 31)	54%	60%	55%
BIA Body Fat %	25.7	26.3	26.0
Estimated Yearly Household Income (\$)			
7500 or less	4%	9%	5%
7501 to 15000	23%	0%	19%
15001 to 25000	34%	9%	29%
25001 to 35000	11%	27%	14%
35001 to 50000	4%	55%	14%
50001 to 75000	4%	0%	4%
75001 or higher	9%	0%	7%
Don't Know	11%	0%	9%

Measures

Body Mass Index (BMI). BMI is one of the most commonly used methods to assess whether a person is overweight or obese (Prentice & Jebb, 2001). In order to calculate the BMI, one divides the person's weight in kilograms by the square of their height in meters. This value is then compared to age and gender-specific cutoffs in order to determine the person's comparative health. When measurements are obtained correctly, BMI has been supported as a reliable and valid indicator of overweight and obesity for clinical, screening, and surveillance purposes and thus, has been utilized by the CDC as the criteria for which children and adolescents are categorized as overweight or obese (CDC, 2011; Himes, 2009). Additionally, BMI has demonstrated criterion validity in its association with several health outcomes (see Wabitch, 2000a; 200b). The reliability of BMI has been found to be generally high compared to other methods of body composition measurement (e.g., Skinfold thickness; Dietz & Bellizi, 1999).

As suggested by Himes (2009), evaluators were trained in measuring height and weight in order to reduce measurement error. Weight was measured using electric scales, and height was measured using stadiometers. Students were instructed to remove their socks and shoes, and to stand up straight on the stadiometer with their back up against the vertical measuring stand. A sliding horizontal piece was moved down to the top of the participant's head and the indicated height in centimeters was recorded. This process was completed twice for each participant in succession and values were averaged to reduce measurement error. Participants were then instructed to stand still with their hands at their sides on electronic scales until weight (kg) was displayed on the screen. Children were not provided feedback on their weight status. Values of height and weight were recorded

immediately on site and BMI was later calculated by dividing weight (kg) by the square of their height (m²). BMI values were compared to age and gender norms in order to determine weight status.

Bioelectrical Impedance Analysis (BIA). BIA determines the electrical impedance (i.e., opposition to flow of electric current) through body tissue by capitalizing on the concept that fat-free mass contains nearly all of the body's conducting electrolytes (Tyrell et al., 2001). Biometric data such as height, sex, and age are used along with impedance information in a prediction equation in order to provide estimates of fat-free mass, total body water and body fat (Tyrell et al., 2001). Gutin et al. (1996) found high internal consistency and test-retest reliability for foot-to-foot BIA (ICC > .99; 2% or less change from trial 1 to trial 2). Body fat measured by foot-to-foot BIA has also been found to have a very high correlation with a criterion measure of DXA, indicating convergent validity ($r = .98$; Tyrell et al., 2001). There is some evidence indicating that foot-to-foot BIA overestimates fat mass and body fat and underestimates fat-free mass (Tyrell et al., 2001). Although Buchholz et al. (2004) caution against the use of BIA in single measurements (i.e. screening individuals) due to increased chance of error, they acknowledged that BIA may be acceptable for determining differences in body composition between groups.

BIA was measured using electric scales. Participants were instructed to remove their shoes and socks, step on the scale with their feet touching the metal nodes and stand still until the reading was complete. Digital values were displayed on the scale within moments and the participant's percent body fat was recorded. Occasionally, errors in readings required this process to be repeated due to a number of causes (e.g., the child

moving or feet not in position). Scales were cleaned using disinfecting wipes between each participant's assessment.

ActiGraph GT1M Accelerometer. Accelerometers are electronic devices attached to the body (typically the hip or lower back) in order to provide quantitative information regarding body accelerations at specified time intervals, called epochs. Accelerometers provide a direct measure of movement and are able to measure the intensity of movements at a specified time period. Additionally, triaxial accelerometers (essentially three uniaxial accelerometers combined in one device) are able to provide information on movements along three planes (i.e., up-and-down, side-to-side, forward-and-backward) instead of merely up-and-down motion. Accelerometer data is recorded in "counts" over a given period of time, which are often entered into equations to yield more interpretable measures of physical activity such as MVPA or energy expenditure (Troiano, 2006). The use of accelerometers as a measure of physical activity has been investigated in a number of studies with the majority of research showing strong correlations between accelerometer measurements and energy expenditure or exercise intensity as measured by heart rate telemetry, DLW, and behavioral observation (Freedson, et al., 2005; Janz, 1994; Trost, et al., 2005). The use of triaxial accelerometers to assess energy expenditure in children has also been validated in simulated free-living conditions such as sitting, writing, laying down, cycling, stepping, jogging, and performing tasks related to basketball, soccer, and tennis (Sun et al., 2008). The ability of accelerometers to record and store data on the device allows for measurements to be taken for days at a time, allowing for the estimation of typical daily or weekly MVPA time. In order to appropriately estimate energy expenditure or activity intensity using

accelerometers, great care must be taken in selecting the product, placing the device on and preparing participants, setting appropriate time intervals (epochs), allowing a sufficient number of monitoring days, and setting appropriate cut points for translating counts into physical activity or energy expenditure units (Trost et al., 2005).

The ActiGraph accelerometer is small, lightweight, and the most widely used monitor (ActiGraph LLC, Fort Walton Beach, FL; Trost, 2007). Advantages of the ActiGraph GT1M include direct USB connection, 1Mb unit memory, self-calibrating digital accelerometer, and documented validity in children and adolescents (Freedson et al., 2005; Trost et al., 2005). Reliability of the ActiGraph accelerometer has been found to depend on the number of days monitored, from moderate in one day (ICC = .45) to high after eight days (ICC = .90; Garnier & Benefice, 2006). The ActiGraph Accelerometer has also been validated by direct observation (Hands et al., 2006), DLW, and other accelerometers (see De Vries et al., 2009).

Previous studies indicate that due to the intermittent, erratic activity often found in children and adolescents, shorter epochs (i.e., 5 seconds or less) over longer periods of time are suggested in order to capture physical activity (McClain et al., 2008). Thus, in the current study recordings were made in 3s epochs and students were instructed to wear the device over course of 5 days as much as possible except while sleeping or in water such as bathing or swimming. Users were instructed to attach the device on the hip at the mid-axillary line and shown how to fasten clasps on provided, appropriately-sized belts. After explaining and demonstrating the attachment and placement of the device, participants were asked to place the device on themselves and were reinstructed if necessary.

Progressive Aerobic Cardiovascular Endurance Run (PACER) and Healthy Fitness Zone (HFZ). The PACER has been developed as part of the FITNESSGRAM battery of physiological assessments in order to investigate aerobic fitness (Cooper Institute for Aerobics Research, 1999). The PACER is a multistage 20-meter shuttle run that progresses from easy to difficult over time. Participants aim to complete as many laps back and forth across the 20-meter distance as many times as possible during the shuttle run and are eliminated if they do not cross the appropriate distance within the allotted time for the second time. Students are prompted with voice commands, beeps that signify the start and end of each lap, as well as “triple beeps” that indicate that a minute has passed and the pace will increase. Equations have been developed that take into account age and PACER results to predict $VO_2\text{max}$, which is an ideal measure of aerobic capacity (Leger, Mercier, Gadoury, & Lambert, 1988). Several considerations must be made when administering and interpreting PACER data. A repeated-measures study performed with children and adolescents in grades 4 to 8 indicated that children tend to improve their PACER scores over the school year with drops over the summer (Butterfield, Lehnhard, Mason, & McCormick, 2008). Authors hypothesized that the lack of forced activity over the summer may have accounted for their drop in performance but did not collect data on activities performed in the summer months. Authors also observed practice effects with improved pacing on the test over time, which could be related to increased motivation as students seemed interested in improving their scores at each administration (Butterfield, et al., 2008). Despite these setbacks, the PACER has been validated and successfully utilized to investigate group differences in cardiovascular fitness in large-scale research studies (Beets, Pitetti, & Cardinal, 2005; Chun, 2000). The

PACER is an attractive measure of physical fitness as multiple students can be measured simultaneously in a relatively brief period of time. Furthermore, there is a relatively low cost of required materials (e.g., pre-recorded beep prompts, cones, and measuring tape), which can be easily transported to different observation sites. Even though the PACER requires participants to perform physical activity in a group setting, Gao (2008) found that enjoyment of physical education did not contribute significantly to PACER performance, and perceived competence accounted for less than 20% of the variance in PACER performance. Changes in PACER performance is impacted by maturation through natural changes to aerobic capacity and running economy (Cooper Institute for Aerobics Research). For example, girls have a natural decrease in aerobic capacity after age 10 that is offset by improved running economy through adolescence. Boys; however, tend to have progressive improvements in scores through adolescence due to improved running economy and constant relative aerobic capacity. Thus, norm-based standards are utilized to determine whether PACER scores are indicative of good overall fitness using age and gender to determine whether a score falls within the Healthy Fitness Zone (HFZ).

Procedures

Data were collected as part of a larger study funded by the Robert Wood Johnson Foundation. The original study evaluated the implementation of an evidence-based physical activity instruction curricula (SPARK; *Sports, Play, and Active Recreation for Kids*) in comparison to standard physical activity curricula in a naturalistic Out-Of-School (OOS) setting (see Thaw et al., 2014). Data in this study were collected between March and September of 2010 in Miami-Dade County, Florida. Consents were reviewed and signed by parents while a separate assent form was provided for children and

adolescents. These forms were available in both English as well as Spanish, and bilingual staff members were present to explain details in the parent's preferred language.

Participant's parents were offered incentives of \$25 gift cards to Target for volunteering to participate in the study and received their reward regardless of compliance.

Once consent was established, participants removed their footwear and had their height assessed using stadiometers, and weight/BMI measured using electronic scales (see descriptions above). Participants were given their accelerometers one at a time, instructed on how to wear the device properly, and checked to ensure they learned how and where to place the device on their own. Administrators used measuring tape to mark 20-meter lanes using small brightly-colored cones. Areas of administration varied from site to site but efforts were made to set aside the administration area in flat, clear areas away from other students and free from distraction. Participants were given instructions on how to complete the PACER and briefly quizzed to ensure their knowledge. Once participants were ready, a pre-recorded FITNESSGRAM 20-meter PACER CD was played via portable stereo and pre-recorded instructions were repeated before the test began. Music and verbal prompts provided feedback on the increasing pace of the test. After at least five days of accelerometer wear-time, staff members returned to the site, collected accelerometer devices from participants, provided gift card rewards, and data was downloaded from devices to be analyzed.

Analyses

Data was validated, reduced, and translated into accelerometer counts, and time spent engaging in physical activity using ActiLife 5 data analysis software. Other analyses were performed using Predictive Analytic SoftWare (PASW) Statistics 18.0 software. Primary analyses include a hierarchical regression to determine the relative influence of demographic variables, body composition, and physical activity on physical fitness. In the first block, gender (male, female), income level (from annual income of \$7,500 or less to more than \$75,000 per year), race/ethnic category (Hispanic, Non-Hispanic Black), and age (months) were entered as variables predicting physical fitness as measured by PACER laps. In the second block, body composition variables were added to the model including percentile of BMI and body fat percentage (as measured by foot-to-foot BIA). In the final model, accelerometer counts per minute were added to the regression. This procedure was repeated as a logistic regression with the same predictor variables but with the outcome variable as the dichotomous variable of performance on the PACER falling in the Healthy Fitness Zone or Not Healthy Fitness Zone. Descriptive statistics will be presented along with any between-group differences in ethnic-gender subgroups. Two-way ANOVAs will examine differences between gender and race/ethnic group on times spent in physical activity categories. Correlations between primary variables are also be presented.

Hypotheses

A priori hypotheses were made such that demographic variables (gender, income, race, and age) will significantly predict fitness (PACER laps) at the first block. When added to the model, body composition variables (BMI and BIA-measured body fat

percentage) will significantly predict fitness over and above demographic variables. Then, physical activity (accelerometer counts) will significantly contribute to predicting fitness above and beyond the cumulative impact of demographic and body composition variables when added to the final model. In the logistic regression, lower body fat, lower BMI, male gender, younger age, and higher accelerometer counts will be associated with higher likelihood of HFZ status. A priori hypotheses were made such that boys will have greater MVPA time and PACER performance compared to girls regardless of race/ethnic category. Racial/ethnic category will interact with gender such that the Non-Hispanic Black male group will have significantly more time spent in moderate-to-vigorous physical activity.

Statistical Methods

Accelerometry data. Accelerometer data was downloaded, reduced, and converted using ActiGraph ActiLife 5 software as well as Santech MeterPlus 4.3 software. A day of valid wear time was determined using the 70/80 rule, which determines the amount of time that 70% of the sample has recorded accelerometer data (approximately 9.24 hours in this study), and requires 80% of this period as the minimum amount of time in a valid day (approximately 7.38 hours in this study; see Ward, Evenson, Vaughn, Rodgers, & Troiano, 2005). To be conservative in ensuring use of valid data, a minimum of 8 hours of wear-time was required in order to be considered a valid day of activity measurement in this study. A reading of zero for 60 or more consecutive minutes was considered to be non-wear time and was excluded from the analysis. Participants provided an average of 4.53 (SD = 1.77) valid days of data, with a mean of 12.78 (SD = 2.17) valid hours of accelerometer data per valid day. Ethnic/gender

subgroups did not significantly differ on valid days or valid hours (see Table 6). There were no significant differences in activity intensity on weekdays compared to weekends (see Table 7). Accelerometer measurements were considered in counts per minute to provide a continuous estimate of physical activity that could later be translated into other units (e.g., physical activity categories, kilocalories, METs).

Statistical Analysis. The skewness and kurtosis of all variables were examined to ensure a normal distribution before being utilized in the study. Bivariate correlations were examined between key variables using Pearson's correlations. The primary model of interest in the study is a hierarchical regression with three blocks of variables utilized to predict fitness as measured by PACER laps. In the first block, demographic variables were entered, including gender (0 = male, 1 = female), race/ethnic category (Hispanic or Non-Hispanic Black), household income estimated by the child's parent (grouped from less than \$7500 to over \$75000), and age at the time of evaluation. In the second block, body composition variables were added including BMI percentiles calculated from height and weight measurements, and body fat percentage measured by foot-to-foot BIA. In the final block, physical activity (counts/min) was added to the model. An alpha level of 0.05 was used in all significance testing unless otherwise noted. Statistical analyses were conducted using Predictive Analytic SoftWare (PASW) Statistics, version 18.0

Table 6*Descriptive characteristics*

	Hispanic (<i>n</i> = 47)				Non-Hispanic Black (<i>n</i> = 11)			
	Male (<i>n</i> = 21)		Female (<i>n</i> = 26)		Male (<i>n</i> = 6)		Female (<i>n</i> = 5)	
	M	SD	M	SD	M	SD	M	SD
Age (years)	10.10	2.86	8.96	2.55	9.00	2.00	10.00	1.58
BMI (kg/m ²)	21.92	4.76	20.73	4.26	21.97	4.92	21.56	2.81
BMI Percentile for Age	81.96	23.00	78.38	25.20	81.96	20.60	88.56	5.90
BIA (% body fat)	23.26	8.17	27.46	8.29	26.97	8.83	28.08	4.07
Activity (counts/min)	459.18	159.97	368.11	140.25	503.67	97.43	428.79	122.25
Cardiovascular Fitness (PACER Laps)	23.81	17.06	14.85	6.76	18.50	14.22	16.00	5.70
Valid Accelerometer Days	4.67	1.71	4.69	1.78	3.83	2.04	4.00	1.87
Valid Accelerometer Hours per Valid Day	12.53	1.73	12.90	1.84	14.29	4.13	11.43	1.90

Table 7*Percentage of Time Spent in Physical Activity Categories on Weekdays and Weekends**

	Sedentary	Light	Moderate	Vigorous
Weekday Only	83.97	11.65	3.96	0.41
Weekend Only	83.78	12	3.79	0.43

*Using cut-points developed by Puyau et al. (2002)

CHAPTER III: RESULTS

Descriptive Statistics

Descriptive statistics in the sample by race/gender subgroups are presented in Table 6. Overall, the sample consisted of approximately 81% Hispanic participants ($n = 47$) and 19% Non-Hispanic Black participants ($n = 11$). Additionally, 46% were boys ($n = 27$) while 54% were girls ($n = 31$). The average BMI was near the 81st percentile, based on expected BMI for age (CDC, 2011). The mean body fat percentages were approximately 24.1% for boys and 27.6% for girls, with 25.6% for Hispanic participants and 27.5% for Non-Hispanic Black participants (see Table 6 for subgroup statistics). An independent-samples t -test indicated that males ran significantly more PACER laps ($M = 22.6$, $SD = 16.4$) compared to females ($M = 15.0$, $SD = 6.5$), $t(56) = 2.39$, $p = .021$, $d = .63$. Thus, on average, males ran approximately 7.6 more laps compared to their female counterparts. Males also had statistically significantly higher accelerometer-measured counts per minute ($M = 469.1$, $SD = 147.9$) compared to females ($M = 377.9$, $SD = 107.5$), $t(56) = 2.709$, $p = .008$, $d = .72$. This is can be interpreted as a medium effect size, according to Cohen (1988).

Percentage of time in physical activity categories were calculated based on cut-off values using accelerometer counts as suggested by Puyau, Adolph, Vohra, and Butte (2002). This particular equation was chosen as it was developed and validated using children of an age similar to the current sample (6-16 years). Cut points were set using the following categories: Sedentary (<800 counts/min), Light (<3200 counts/min), Moderate (<8200 counts/min), and Vigorous (≥ 8200 counts/min). The average number of minutes per day spent in each physical activity category can be found in Table 8. Overall,

less than three minutes per day were spent engaging in vigorous physical activity, and approximately 30 minutes per day were spent engaging in moderate physical activity. Since differences in validated accelerometer wear-time may confound comparisons between ethnic/gender subgroups (see Table 6), percent of validated wear-time spent engaging in varied physical activity intensities are presented in Table 8. Overall, the overwhelming majority of time spent wearing the accelerometer was in sedentary activity (approximately 84%), including on weekends (see Table 7). Overall, less than one-half percent of time was spent in vigorous physical activity (see Table 8 and Figure 3). A two-way ANOVA of percent of time spent in each physical activity level revealed statistically significant main effects for gender and race/ethnic category at every activity level with the exception that there was no statistically significant main effect of gender on vigorous physical activity time. Results indicate that females spent a greater percentage of time in the sedentary category; while males had a greater proportion of time in light and moderate activity (see Table 9 and Table 10). Excluding the vigorous category, between 6% and 16% of the difference in physical activity percentage could be accounted for by gender. Furthermore, Hispanic participants spent a greater proportion of time in the sedentary category, while Non-Hispanic Black participants spent a greater proportion of time in light, moderate, and vigorous activity categories (see Table 9 and Table 10). Between 8% and 16% of the difference in physical activity could be attributed to race/ethnic category. There were no significant interaction effects between gender and race/ethnicity on percentage of physical activity time, meaning that the effects of gender and race/ethnicity on physical activity time do not depend on one another (see Table 9).

Table 8*Daily Minutes in Physical Activity Categories**

	All participants (N = 58)	Hispanic (n = 47)		Non-Hispanic Black (n = 11)	
		Male (n = 21)	Female (n = 26)	Male (n = 6)	Female (n = 5)
Sedentary	641.89	627.43	663.41	661.43	567.24
Light	92.31	90.23	84.54	138.64	85.89
Moderate	29.75	31.25	23.41	52.48	29.13
Vigorous	2.99	2.8	2.61	4.82	3.52

*Using cut-points developed by Puyau et al. (2002)

Table 9*Percent of Validated Time in Physical Activity Categories**

	Hispanic (n = 47)				Non-Hispanic Black (n = 11)			
	Male (n = 21)		Female (n = 26)		Male (n = 6)		Female (n = 5)	
	M	SD	M	SD	M	SD	M	SD
Sedentary	83.53	5.70	85.67	3.37	77.67	5.30	83.19	4.45
Light	11.96	4.10	10.97	2.51	15.56	3.62	12.21	3.33
Moderate	4.15	1.73	3.02	0.98	6.18	2.15	4.10	1.27
Vigorous	0.37	0.23	0.33	0.24	0.60	0.39	0.50	0.19

*Using cut-points developed by Puyau et al. (2002)

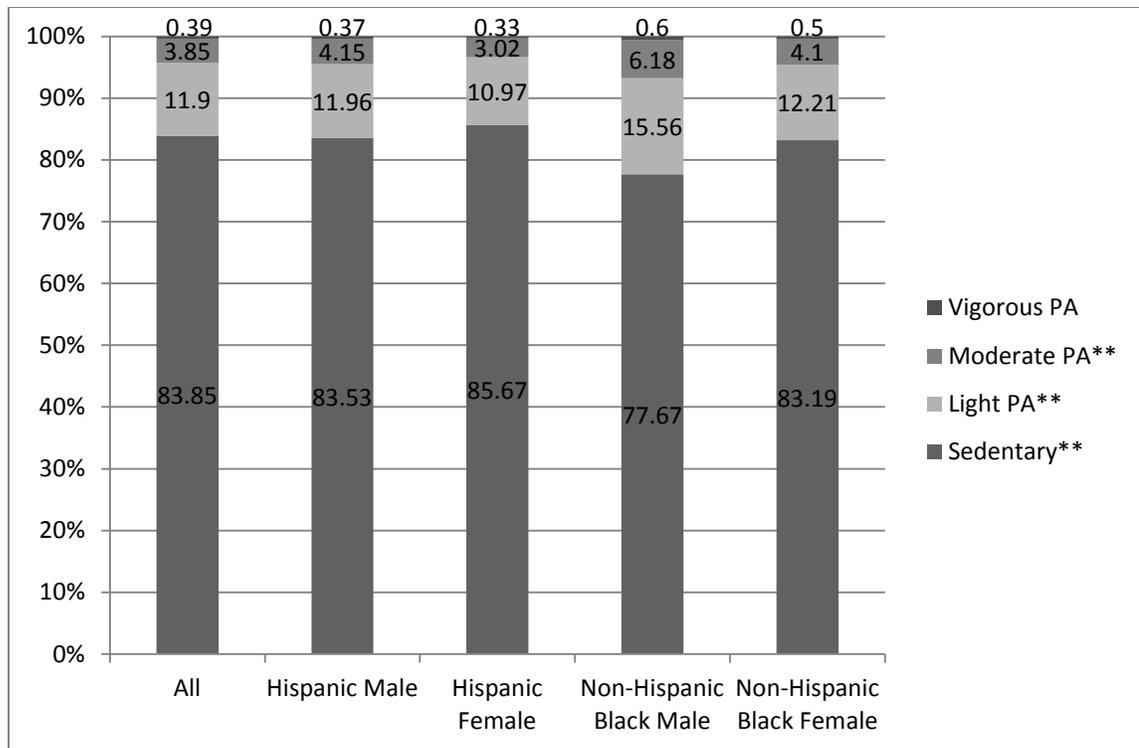


Figure 3. Percent of time engaging in different intensities of physical activity by ethnic-gender subgroup. Main effects of gender and race/ethnic category were present for sedentary time, light physical activity, and moderate physical activity.

Table 10

Two-Way Analysis of Variance of Physical Activity by Gender and Race/Ethnicity

Activity level	Gender			Ethnicity			Gender*Ethnicity		
	F	p	η^2	F	p	η^2	F	p	η^2
Sedentary	6.07	0.02	0.10	7.18	0.01	0.12	1.18	0.28	0.02
Light	3.70	0.06	0.06	4.63	0.04	0.08	1.10	0.30	0.02
Moderate	10.77	<.01	0.16	10.18	<.01	0.16	0.96	0.33	0.02
Vigorous	0.70	0.41	0.01	5.27	0.03	0.09	0.12	0.73	0.00

Note. *df* = 1, 57

Correlational Analysis

Bivariate correlations of primary variables are presented in Table 11. Physical activity, as measured by accelerometer, was not significantly correlated with age, BMI percentile for age, percent body fat measured by foot-to-foot BIA, or physical fitness as measured by PACER laps. BMI percentile was negatively correlated with PACER laps and strongly positively correlated with percent body fat. Age was also strongly correlated with PACER laps, which indicates that older youth complete more laps. Percent body fat as measured by BIA was also had a significant negative correlation with PACER laps, which suggests that those with more body fat complete fewer PACER laps.

Table 11
Bivariate Correlations

	BMI Percentile	BIA	Counts/min	PACER Laps
Age	.043	0.146	-0.181	.658*
BMI Percentile		.776*	-0.003	-.260*
BIA			-0.203	-.337*
Count/min				.022

*Significant at the .05 alpha level

Regression Analysis

A hierarchical regression model was utilized to predict PACER laps with three distinct blocks. The hierarchical regression model allows one to determine the amount of variance attributed to the model above and beyond variables in the previous block. The first block of demographic variables (gender, income, race/ethnic category and age) contributed significantly to the regression model, $F(4, 53) = 12.48, p = <.01$, and accounted for 49% of the variation in PACER laps. Introducing body composition variables (BMI percentile and BIA-measured body fat) significantly contributed an

additional 16% of the variation in PACER laps, $F(2, 51) = 11.24, p < .01$. However, adding physical activity (counts/min) to the model did not contribute over and above previously entered variables, $F(1, 50) = 0.191, p = .66$, and only accounted for an additional 0.1% of the variance of PACER laps.

The unique contribution of each variable while considering the impact of other variables can be found in Table 12. Here, age and BIA-measured body fat were the only significant contributors to PACER laps while considering other variables. These results indicate that each additional year of age contributed to an increase in 3.48 PACER laps. The impact of age seemed to dominate the final model, accounting for nearly 45% the variance of PACER laps, even while considering the influence of other variables. Additionally, BIA-measured body fat contributed to PACER performance such that, for every one percent increase in body fat, participants would run .75 fewer PACER laps.

Table 12

Hierarchical Regression Variables Contributing to PACER Laps

Variable	B	SE	<i>t</i>	<i>p</i>	sr ²
Gender	-2.25	2.48	-0.91	0.37	0.01
Income	0.39	0.57	0.69	0.49	0.00
Race/Ethnicity	-1.17	2.77	-0.42	0.68	0.00
Age*	0.29	0.04	7.94	<.01	0.45
BMI Percentile	0.04	0.08	0.50	0.62	0.00
BIA*	-0.75	0.24	-3.10	<.01	0.07
Counts/min	0.00	0.01	0.44	0.66	0.00

*Statistically significant at the .05 alpha level

Since the primary aim of the current study was to examine the relative contribution of physical activity above and beyond demographic and body composition variables, the entire process was repeated with age excluded from the model to more closely examine such relationships without the overwhelming influence of age. In the

repeated regression model, the initial block (with age excluded) accounted for 9.8% of the variance in PACER laps, which was not statistically significant, $F(3, 54) = 1.95, p = .13$. In the second block, the addition of BMI percentile and BIA added 8.8% of the variation in PACER laps; however, this was also not statistically significant at the .05 alpha level, $F(2, 52) = 2.81, p = .07$. Finally, accelerometer-measured counts per minute in the final block contributed 0.9% of the variation in PACER laps, which was not a statistically significant contribution, $F(1, 51) = .57, p = .45$. Individual predictors in the final model are presented in Table 13. Here, gender was the only statistically significant variable, contributing approximately 7% of the variance, when considering the impact of all other variables in the model. This result suggests that being a male led to an increase in approximately 7.5 PACER laps, while controlling for demographic and body composition factors (without considering the impact of age).

Table 13

Hierarchical Regression Variables Contributing to PACER Laps (Age Excluded)

Variable	β	SE	t	p	sr^2
Gender*	-7.51	3.56	-2.11	0.04	0.07
Income	0.48	0.85	0.57	0.57	0.01
Race/Ethnicity	-0.81	4.13	-0.20	0.85	0.00
BMI Percentile	-0.06	0.12	-0.49	0.63	0.00
BIA	-0.36	0.35	-1.01	0.32	0.02
Counts/min	-0.01	0.01	-0.76	0.45	0.01

*Statistically significant at the .05 alpha level

In order to further investigate the role of physical activity in contributing to cardiorespiratory fitness, each participant's PACER performance was categorized as falling in the Healthy Fitness Zone (HFZ) or Not in Healthy Fitness Zone (NHFZ), based on standards provided by FitnessGram and The Cooper Institute (See Plowman &

Meredith, 2013). This standard is developed using norm-based data which indicates the level of performance at which a child may or may not be at risk for the development of health issues. Descriptive statistics of those classified in the HFZ and NHFZ are presented in Table 14. Without considering the impact of other factors, those classified in the HFZ were significantly younger, had a lower BMI (though, not a statistically significantly lower BMI percentile), and had less body fat as measured by BIA. Regarding counts/min, there was no significant difference between groups, with a small effect size (Cohen's $d = .33$). The percent of males and females falling in the HFZ and NHFZ did not vary by gender, $X^2(2, N = 58) = 3.29, p = .071, r = .24$.

The dichotomous variable of HFZ classification was used as the dependent variable in a logistic regression using all relevant variables, including age. Results indicated that all variables included in the model correctly identified HFZ performance in 87.9% of the sample in the final model. Odds ratios and p -values of each variable in the final model are presented in Table 15. Overall, gender, age, and body fat (BIA) were significant predictors of HFZ classification. Results indicate that, holding other factors constant, females are more than 25 times more likely to be classified in the Healthy Fitness Zone when controlling for other variables in the model. Furthermore, older participants and those with higher percent body fat were less likely to be placed in the HFZ. Physical activity (counts/min) was not significantly associated with HFZ when considering other variables. In order to further investigate the relative contribution of physical activity, the same regression was executed with gender and age removed from the model. Odds ratios and p -values of this model are presented in Table 16. Here, none

of the predicting variables were unique contributors in predicting HFZ status at the .05 alpha level.

Table 14
HFZ Descriptive Statistics

	In HFZ (<i>n</i> = 43)		In NHFZ (<i>n</i> = 15)	
	M	SE	M	SE
Age (months)*	108.9	4.15	147.4	6.18
BMI*	20.12	0.58	24.78	1.06
BMI Percentile	78.59	23.19	87.61	20.29
BIA*	24.41	1.22	30.32	1.78
Counts/min	432.05	21.273	386.78	30.5

*Significant at .05 alpha level

Table 15
Logistic Regression Variables Predicting HFZ Categorization

	β	SE	<i>P</i>	OR
Gender*	3.252	1.403	.020	25.641
Race/Ethnicity	.104	1.263	.934	1.109
Income	.070	.211	.741	1.072
Age (months)*	-.058	.020	.004	1.060
BMI Percentile	.068	.039	.082	1.070
BIA*	-.391	.152	.010	1.479
Counts/min	.004	.004	.361	1.004

*Significant at .05 alpha level

Table 16
Logistic Regression Variables Predicting HFZ Categorization (Age and Gender Excluded)

	β	SE	<i>p</i>	OR
Race/Ethnicity	0.045	0.829	0.956	1.046
Income	-0.059	0.167	0.724	1.06
BMI Percentile	0.021	0.027	0.431	1.021
BIA body fat	-0.148	0.08	0.065	1.16
Counts/min	0.001	0.003	0.674	1.001
Constant	3.098	2.17	0.153	22.143

*Significant at .05 alpha level

CHAPTER IV: DISCUSSION

Implications of Findings

The primary aim of the current study is to explore the relative utility of demographic variables, body composition, and physical activity as predictors of fitness status in an ethnically diverse group of children and adolescents. Results suggest a fairly strong and consistent association between certain demographic variables (age, gender) and body composition (body fat) with “fitness,” as measured by 20m shuttle run (i.e., the “beep test”). However, there was a notable lack of association of accelerometer-measured physical activity with fitness. This result is contrary to previous findings (e.g., Aires et al., 2010; Lohman et al., 2008). One aspect of the current study was determining the empirical value of measures of physical activity and fitness which fall on a continuum of psychometric properties (i.e., reliability, validity), and clinical utility. Accelerometer-measured physical activity is promising as a measure given its high clinical utility when compared to other measures such as expensive, lab-collected DLW; however, the current study did not find a strong association between accelerometer-measured counts and other indicators of health status, such as BMI, BIA-measured body fat, and physical fitness. These results therefore run contrary to previous findings where physical activity was negatively associated with fat mass (Rennie et al., 2005). While feasibility in use of accelerometers in measuring large groups is a strength, costs total quickly (\$200-\$300 per unit), there is risk of damaged or lost units, and accessing comprehensible results necessitates time, expensive software (nearly \$1700 for single license), and high expertise to interpret results appropriately. Given these barriers along with the surprising results in the current study, perhaps other avenues of measuring physical activity in

groups should be investigated as developments in technology continue to advance (see Future Research section below).

The lack of association between physical activity and fitness in this study may indicate that resources aimed at improving fitness in children and adolescents could be best utilized to target other important factors such as body fat, rather than physical activity. This may seem paradoxical as increased physical activity is sometimes utilized as an intervention for lowering body fat. However, physical activity is only one factor in lowering body fat and; in fact, some research suggests only time spent in vigorous physical activity is associated with lower body fat when considering demographic variables (Gutin et al., 2005). Furthermore, there is growing evidence in the literature suggesting that vigorous (as opposed to moderate) physical activity is necessary to generate strong associations with cardiorespiratory fitness in youth (Gutin et al., 2005; Martinez-Gomez et al., 2010; Ruiz et al., 2006; Taber et al., 2014). Participants in the current study spent very little time in vigorous physical activity (about 3 minutes per day), which may be one reason why physical activity was not found to be associated with fitness. While outside the scope of the current study, diets with high caloric intake, as well as diets high in fat likely contribute to high body fat and obesity (e.g., Lamarche, 1993; Sacks, Bray, & Carey, 2009). Thus, interventions aimed at long-term adherence to low-fat diets which limit caloric intake may prove beneficial in improving cardiovascular fitness in youth.

It is also important to note that the lack of association between physical activity and physical fitness found in the current study does not necessarily negate the positive impact that physical activity may have on other indicators of health and overall. Overall,

results indicate a dire need to increase physical activity time and reduce body fat regardless of demographic category, with specific targeted interventions aimed at promoting increased physical activity in Hispanic and female groups. When considering overall fitness; however, males are at particular risk for being classified as “unfit,” despite their higher physical activity and superior performance before consideration of norm-based standards. Indeed, boys may not be receiving as much focus in interventions to improve fitness since they seem to be more active and performing better than their female counterparts, overall.

Interpretation of Findings by Analysis

BMI and BIA-measured body fat. Generally, body size and composition information provided by the final sample reflects the continued necessity for interventions to improve health status in ethnically diverse youth. Youth in this sample average placement was at the 81st percentile of BMI compared to similarly-aged peers of the same gender. This indicates that the average youth included in the study nearly met criteria for overweight status as measured by BMI (85th percentile; CDC, 2011). With 24.1% and 27.6% body fat, respectively, both males and females fell near the 75th percentile of body fat compared to percentiles of body fat based on NHANES data (Laurson, Eisenmann, & Welk, 2011). Taken in tandem, these results seem to indicate that the youth in the sample are relatively heavy and have a high body fat percentage compared to similarly aged peers in the U.S. Comparisons of body fat between genders in children can be challenging due to natural changes that may occur due to maturation. For example, body fat in boys tends to decrease in childhood then increase in adolescence, while body fat tends to only increase in girls as they age, and at a higher overall

percentage (Ogden, Li, Freedman, Borrud, & Freedman, 2011). While there may be a number of sociocultural factors that might contribute to differences in body fat between ethnic subgroups (e.g., Dodd, Briefel, Cabili, Wilson, & Crepinsek, 2013; Onge, Jarron, & Kreuger, 2011), the current study failed to yield statistically significant differences in BIA-measured body fat between Non-Hispanic Black participants ($M = 27.5\%$) and Hispanic participants ($M = 25.6\%$).

Physical activity and ethnic/gender subgroups. Based on accelerometer-measured activity time, youth in the current study were sedentary for the vast majority of their wear time and spent very little time engaging in vigorous activity. Given an average of just under 13 hours of wear time, youth spent approximately three minutes per day engaging in vigorous activity. Thus, vigorous physical activity time in the current study is lower than previous findings, such as Troiano et al., 2008, (i.e., 2-5 minutes per day vs. 5-8 minutes per day). Approximately 30 minutes per day were spent engaging in moderate physical activity, which is half the amount of the recommended moderate-to-vigorous physical activity (MVPA) recommended for youth by the World Health Organization (WHO, 2011). Boys in this sample spent approximately 12 more minutes per day in moderately intense physical activities and had approximately 13 minutes per day less sedentary time compared to girls; a finding which is consistent with large-scale studies (see Troiano et al., 2008). Non-Hispanic Black males seemed to be the most physically active while Hispanic females were the most sedentary. Non-Hispanic Black males had nearly twice the daily MVPA time (57 minutes) compared to Hispanic females (26 minutes). Previous research indicates Black female youth as demonstrating lower physical activity and fitness levels compared to other ethnic groups (Lohman et al.,

2008); results of the current study expand on these findings and suggest that the physical activity and fitness levels of Hispanic females are similar; if not lower, than for Black females. Belcher et al. (2010) found that Non-Hispanic Black youth recorded higher MVPA compared to Non-Hispanic White and Mexican American youth. Similarly, Black participants in the current study showed greater physical activity time as well as less sedentary time compared to their Hispanic counterparts. However, further research is necessary to determine the factors that contribute to these differences between race/ethnic subgroups in physical activity. For example, there is evidence suggesting that there may be cultural preferences and norms which contribute to such differences. For example, using data from the National Health Interview Survey ($N = 17,455$), Onge, Jarron, and Kreuger (2011) found that Non-Hispanic Black groups are more likely to participate in fitness activities such as running and weight lifting, while Mexican American groups are more likely to engage in team sports.

Correlational analysis. The strongest correlation between primary variables was between BMI Percentile and BIA-measured body fat percentage ($r = .776$). BMI has been criticized for its relatively poor association with body composition (e.g., Wheeler & Twist, 2010). However, this study indicates a strong association between BMI and body fat percentage. Since BMI does not directly measure body fat, it is likely a poor indicator of health status for individuals with unique body composition for their height and weight, such as body-builders or elderly populations (Wheeler & Twist, 2010). Thus, one explanation for this strong association in the current study may be that the youth are far less likely to fall into such a unique category. BMI percentile and BIA were inversely associated with PACER laps, suggesting that those with less healthy body composition

(i.e., heavier and more fat) tend to be less physically fit. This finding is one that is not surprising given associations between body composition and future health outcomes. However, accelerometer-measured physical activity was not significantly correlated with physical fitness; which is an unexpected result that is inconsistent with previous research. For example, Lohman et al. (2008) found a statistically significant correlation between accelerometer-measured physical fitness and cardiorespiratory fitness ($r = .16$) in a diverse sample of eighth grade girls, and Aires et al. (2010) found a statistically significant association between accelerometer-measured physical activity and PACER-measured cardiorespiratory fitness ($r = .28$) in Portuguese middle and high school students.

Regression Analyses. The initial hierarchical model including demographic variables, body composition, and physical activity accounted for nearly two-thirds of the variance in physical fitness measured by PACER laps. In the first block, demographic information alone accounted for a high amount of variance of fitness level, even without consideration of body measurements or physical activity. However, the very high impact of age on PACER laps likely accounted for much of this result. In other words, older children were able to run a greater number of laps, likely a function of variables related to development, such as stride length. Yet, when BMI percentile and BIA were added to the model, they contributed a significant amount of variance above and beyond the effect of demographic variables, which indicates that body composition significantly contributes to cardiovascular fitness above and beyond demographic factors. Furthermore, body fat remained a significant predictor of fitness even when considering all other variables. Thus, those with less body fat had improved fitness when controlling for the impact of all

other variables in the model (i.e., demographic variables, BMI, and accelerometer-measured physical activity).

Physical activity measured via accelerometry had no association with fitness, and the effect size was even smaller when considering the influence of other variables in the model. This is an unexpected and notable result in contrast with previously demonstrated associations between physical activity and physical fitness (see Gutin et al., 2005; Lohman et al., 2008; Aires et al., 2010). Even when the model was repeated with age removed, physical activity contributed less than 1% of the variance of fitness and was not a significant addition to the regression model. In this follow-up analysis, being a male was the only significant predictor of improved performance on the PACER test.

When the analysis was repeated using a logistic regression with HFZ status as a dichotomous outcome, those that fell in the HFZ tended to be younger, weigh less for their height, and have less body fat. These associations are consistent with general correlates of health. However, accelerometer-measured physical activity did not differ significantly between groups. Once the impact of stronger factors like age, BMI, and body fat were taken into account, the already small impact of accelerometer-measured physical activity on HFZ categorization became nearly nonexistent. In this analysis, gender, age, and BIA-measured body fat remained as significant predictors of HFZ status. Younger participants had a much greater likelihood of falling in the HFZ such that, for each year increase in age, participants were twice as likely to fall out of the HFZ when taking into account other factors in the model. Body fat was another significant factor in predicting HFZ membership, such that participants were nearly 1.5 times as likely to fall in out of the HFZ range for every 1% increase in body fat as measured by

BIA. Even though boys in this study ran significantly more laps than girls (see Descriptive Statistics section), when considering the impact of all factors in the logistic regression, being female meant the participant was 25 times as likely to fall in the HFZ. It is possible that females in this study simply performed extraordinarily well compared to males once other important factors in the model are accounted for (e.g., age, body fat). For example, BIA-measured body fat was higher in girls and negatively associated with HFZ status; however, once body fat was controlled for in the model, it seems that females actually performed very well compared to their male counterparts. Surprisingly, BMI percentile was not a significant individual predictor of HFZ status in the final regression model. While BMI and BIA-measured body fat are each individually associated with fitness, they are also highly correlated with one another and the final regression model represents unique contributions of variables while controlling for other factors in the model. Thus, it is possible that the overlap of variance between BMI and body fat in contributing to fitness was utilized by BIA-measured body fat as a unique predictive factor, rather than BMI. Secondly, the average BMI percentile for those falling in the HFZ was 78.6 ($n = 43$), and 87.6 ($n = 15$) for the NHFZ group. These percentiles are fairly high and indicate that this sample was relatively heavy for their age compared to national standards which might restrict the impact of BMI on fitness, especially when taking into account other potent factors.

Limitations

One primary limitation to the current study was the relatively low sample size, particularly when comparing ethnic-gender subgroups. This low sample size reduces power, thus requiring greater effect sizes in order to find statistical significance. Effect

size; however, is not influenced by sample size, and a number of analyses in the current study yielded moderate to high effect sizes. Furthermore, statistically significant findings despite low subgroup size suggest larger and more meaningful differences. For example, comparisons between ethnic and gender groups on physical activity yielded moderate to large effect sizes, leading to statistically significant findings despite low sample size. For the primary analysis of the study; however, investigating the relation of physical activity to fitness, the effect size was very low.

Another limitation of this study involves selection bias; specifically, participants were all enrolled in out-of-school programs which frequently include structured and unstructured physical activity periods (see Thaw et al., 2014). While this selection bias may limit generalization of results, it would likely lead investigators to overestimate the amount of habitual physical activity. As these results indicate that habitual physical activity was a poor predictor of fitness, this effect would likely be attenuated in those that are not enrolled in after school programs that include a period of physical activity. While the high ethnic diversity in this sample fills a gap in the research by providing valuable information regarding the health status and fitness of under-represented groups, it also limits the generalizability of the current findings to large groups. All participants in the final analysis identified as either Non-Hispanic Black or Hispanic White, which does not include Non-Hispanic White groups and may not capture differences which may be present within each racial group (e.g., Caribbean Black vs. African American; Mexican vs. other Latin-American groups). Furthermore, even though income was not highly associated with other important factors in the current study, the majority of participants were from low-income families, and this may have limited the ability to capture the

impact that high income might have on health-related variables. Thus, one must take caution in interpreting results as they may not generalize to populations that were not captured in our sample (e.g., other racial and ethnic groups such as Non-Hispanic White, Asian, and Native American; high SES groups) and caution must be taken when comparing results to similar studies comprised primarily of participants from different sociocultural groups. While the high but restricted ethnic diversity is acknowledged as a limitation regarding external validity, it is also part of what makes the current study a unique contribution to the literature as it provides valuable information about these understudied subgroups.

Some potentially important factors in predicting physical fitness were not included such as diet and engagement in vigorous training, sports, or other organized activities. Including a measure of diet would undoubtedly improve the explanatory power of the regression models; however, measuring diet comes with its own set of challenges (e.g., reliance on self-report, recall bias, observer effects, under-reporting) and was outside of the scope and aims of the larger study from which data was gathered. There is also recent empirical evidence indicating that participation in vigorous sports participation is associated with cardiorespiratory fitness (Taber et al., 2014); however, the vigorous physical activity while participating in such sports should have been captured by the accelerometer data in the current study, unless participants did not adhere to treatment protocol and removed the devices before their participation. While the addition of factors like diet would likely provide a more thorough assessment of factors contributing to youth fitness, the primary aim of the current study was not to maximize the amount of variance accounted for in contributing to fitness, but rather to investigate the relative

importance in consideration of important factors including body composition and physical activity. Furthermore, the amount of variance explained in the regression models was relatively high compared to other similar research. For example, the initial hierarchical model predicting PACER laps accounted for 64% of the variance overall, compared to 22% of the variance accounted for in the final model completed by Lohman et al. (2008). This is likely related to the choice of sample utilized by Lohman et al. (i.e., only eighth grade girls) as the current study included two additional factors which are evidently of great importance in contributing to youth fitness: age and gender. Since these two factors have such a great impact in early blocks of the model, it may have limited the impact of physical activity, which was added in the final block (i.e., above and beyond the contribution of all other factors).

Finally, it is important to note that the current study utilized the PACER as the only outcome measure of cardiorespiratory fitness, which comes with its own strengths and weaknesses as a measure (see Measures section), and serves as a correlative measure of cardiovascular fitness. As discussed by Kazdin (2005), it is unlikely that there is one measure which captures all which might be relevant clinically, and multiple measures are ideal to increase the likelihood of appropriately capturing desired constructs. Using the PACER as the sole outcome measure means that one must strongly consider each methodological and psychometric limitation when interpreting overall results, especially given the low sample size of the current study. For example, even though motivation and perceived competence have been found to be minimally impactful on performance (Gao, 2008), this factor must be taken into consideration when interpreting the overall results of the study. However, a primary aim of the current study is to maximize clinical utility

while maintaining good psychometric properties. The PACER has been validated in evaluating cardiorespiratory fitness in large groups (see Measures section), and utilization of other measures of cardiorespiratory fitness would severely limit feasibility and cost-effectiveness.

Future Research

Replication of this study using interventions to improve physical activity could provide a more comprehensive picture of the role that physical activity plays in contributing to physical fitness. Perhaps baseline habitual physical activity does not predict youth physical fitness, but there is growing evidence that an increase in *vigorous* physical activity could relate to an increase in fitness (Aires et al., 2010; Gutin et al., 2010; Patrick et al. 2004; Taber et al., 2014). Thus, assessments of interventions specifically targeting increased habitual vigorous physical activity (e.g., high-intensity interval training rather than light jogging) would be helpful in clarifying the seemingly nebulous association between physical activity and fitness. Ideally, replication of this study would include a diverse group with a sample size that is large enough to make ethnic-gender comparisons with adequate power. Given the high ethnic diversity in the Miami area, it would also be beneficial to include more detailed information regarding participants' racial and ethnic heritage (e.g., inclusion of Caribbean Black and African American rather than grouping into same category). Additional predictors could be included in order to further evaluate the relative contribution of physical activity to fitness levels including diet, participation in organized athletics, and screen time. Finally, assessment of fitness-related cultural factors as well as interactions between fitness and acculturation would provide information on the root cause of disparities between ethnic-

gender subgroups. One might include factors such as motivation, resources, and perceived barriers (e.g., Eyster et al., 1998; Boyington et al., 2008; Schaefer, Salazar, Bruhn, Saviano, & Boushey, 2009).

Despite the lack of association between accelerometer-measured physical activity and fitness in youth, the need for interventions to improve youth physical activity and fitness remains. Due to rapid advancements in technology, wearable physical activity and fitness-tracking devices have become increasingly affordable, accessible, and popular; though, with varying psychometric properties (Evenson, Goto, & Furberg, 2015). Using wearable devices as a route of measuring physical activity in research would likely facilitate measurement of youth physical activity by simplifying participant requirements and improving the likelihood of compliance and adherence to research protocols. The utility and accuracy of each device should be considered as findings suggest varying validity and reliability depending on model and study methodology (Evenson et al., 2015). Some devices may have improved clinical utility but diminished psychometric properties compared to accelerometers like the ones used in the current study. For example, many devices only measure step counts rather than multi-axial movement, and are worn on the wrist which may attribute extraneous arm movements as body movements. Ferguson et. al (2015) recently investigated the relationship between consumer-grade activity monitors (e.g., Fitbit, Jawbone, Nike Fuelband) and research-grade accelerometers (e.g., BodyMedia SenseWear, ActiGraph GT3X+) in adults and found strong correlations in step counts ($r > .8$) and moderate-to-strong correlations between total daily energy expenditure ($r = .74$ to $.81$) and MVPA ($r = .52$ to $.91$). Furthermore, smart phone devices have been identified as a potentially accurate and

convenient way to measure physical activity and provide a convenient avenue for intervention through mobile applications with progress monitoring and feedback (Bort-Roig, Gilson, Puig-Rivera, Contreras, & Trost, 2014). Mobile devices may also allow for improved self-monitoring of physical activity and tracking of diet, leading to improved interventions to combat obesity (e.g., Turner-McGrievy et al., 2013). As technology continues to develop, these devices and mobile applications will likely become increasingly advanced and accurate in their measurement and represent a promising resource for future research and interventions.

REFERENCES

- Abbott, R. A., & Davies, P. S. W. (2004). Habitual physical activity and physical activity intensity: Their relation to body composition in 5.0-10.5-y-old children. *European Journal of Clinical Nutrition*, *58*(2), 285-291. doi:10.1038/sj.ejcn.1601780
- Aires, L., Silva, P., Silva, G., Santos, M. P., Ribeiro, J. C., & Mota, J. (2010). Intensity of physical activity, cardiorespiratory fitness, and body mass index in youth. *Journal of Physical Activity and Health*, *7*, 54-59.
- Amaral, T.F., Restivo, M.T., Guerra, R.S., Marques, E., Chousal, M.F., & Mota, J. (2010). Accuracy of a digital skinfold system for measuring skinfold thickness and estimating body fat. *British Journal of Nutrition*, *105*, 478-484.
- Anderson, D.E. (2007). Reliability of air displacement plethysmography. *Journal of Strength and Conditioning Research*, *21*(1), 169-172.
- Bailey, R. C., Olson, J., Pepper, S.L., Porszasz, J., Barstow, T.J., & Cooper, D.M. (1995). The level and tempo of children's physical activities: An observational study. *Medicine and Science in Sports and Exercise*, *27*, 1033-1041.
- Bachrach, L.K. (2000). Dual energy X-ray absorptiometry (DEXA) measurements of bone density and body composition: Promise and pitfalls. *Journal of Pediatric Endocrinology and Metabolism*, *13*(2), 983-988.
- Beets, M.W., Patton, M.M., & Edwards, S. (2005). The accuracy of pedometer steps and time during walking in children. *Medicine and Science in Sports and Exercise*, *37*(3), 513-520.
- Beets, M. W., & Pitetti, K. H. (2004). 1-Mile run/walk and body mass index of an ethnically

- diverse sample of youth. *Medicine and Science in Sports and Exercise*, 10, 1796-1803.
- Beets, M. W., Pitetti, K. H., & Cardinal, B. J. (2005). Progressive aerobic cardiovascular endurance run and body mass index among an ethnically diverse sample of 10-15-year-olds. *Research Quarterly for Exercise and Sport*, 76(4), 389-397.
- Belcher, B.R., Berrigan, D., Dodd, K.W., Emken, B.A., Chou, C.P., & Spruijt-Metz, D. (2010). Physical activity in US youth: Effect of race/ethnicity, age, gender, and weight status. *Medicine and Science in Sports and Exercise*, 42(12), 2211-2221.
- Bort-Roig, J., Gilson, N. D., Puig-Ribera, A., Contreras, R. S., & Trost, S. G. (2014). Measuring and influencing physical activity with smartphone technology: A systematic review. *Sports Medicine*, 44(5), 671-86.
- Boyington, J. E. A., Carter-Edwards, L., Piehl, M., Hutson, J., Langdon, D., & McManus, S. (2008). Cultural attitudes toward weight, diet, and physical activity among overweight African American girls. *Preventing Chronic Disease*, 5(2), A36.
- Butterfield, S.A., Lehnhard, R.A., Mason, C.A., & McCormick, R. (2008). Aerobic performance by children in grades 4 to 8: A repeated-measures study. *Perceptual and Motor Skills*, 107(3), 775-90.
- Campbell, D. T. (1957). Factors relevant to the validity of experiments in social settings. *Psychological bulletin*, 54, 297-312.
- Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research*. Chicago: Rand McNally.

Chun, D. M. (2000). Validation of criterion referenced standards for the mile run and progressive aerobic cardiovascular endurance tests. *Research Quarterly*.

American Alliance for Health, Physical Education and Recreation, 71, 125-134.

Centers for Disease Control and Prevention (2009). Recommended community strategies and measurements to prevent obesity in the United States. *Morbidity and Mortality Weekly Report, 58* (RR-7), 1-29.

Centers for Disease Control and Prevention. (2011). About BMI for children and teens.

Retrieved from

http://www.cdc.gov/healthyweight/assessing/bmi/childrens_bmi/about_childrens_bmi.html

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.).

Hillsdale, NJ: Lawrence Earlbaum Associates.

Cooper Institute for Aerobics Research. (1999). *Fitnessgram: Test administration manual*. Champaign, IL: Human Kinetics.

De Vries, S.I., Van Hirtum, H.W., Bakker, I., Hopman-Rock, M., Hirasing, R.A., & Van Mechelen, W. (2009). Validity and reproducibility of motion sensors in youth: A systematic update. *Medicine and Science in Sports and Exercise, 41*(4), 818-827.

Lamarche, B. (1993). Effects of diet and physical activity on adiposity and body fat distribution: Implications for the prevention of cardiovascular disease. *Nutrition Research Reviews, 6*(01), 137-159.

Dodd, A. H., Briefel, R., Cabili, C., Wilson, A., & Crepinsek, M. K. (2013). Disparities in consumption of sugar-sweetened and other beverages by race/ethnicity and obesity status among United States schoolchildren. *Journal of Nutrition*

Education and Behavior, 45(3), 240-249.

doi:<http://dx.doi.org.ezproxylocal.library.nova.edu/10.1016/j.jneb.2012.11.005>

Dietz, W.H., & Bellizzi, M.C. (1999). Introduction: The use of body mass index to assess obesity in children. *The American Journal of Clinical Nutrition*, 70(Suppl.) 123-125.

Eisenmann, J.C., Wickel, E.E., Welk, G.J., & Blair, S.N. (2005). Relationship between adolescent fitness and fatness and cardiovascular disease risk factors in adulthood: The Aerobics Center Longitudinal Study (ACLS). *American Heart Journal*, 149(1), 46-53.

Elberg, J., McDuffie, J.R., Sebring, N.G., Salaita, C., Keil, M., Robotham, D., ... Yanovski, J.A. (2004). Comparison of methods to assess change in children's body composition. *The American Journal of Clinical Nutrition*, 80, 64-69.

Ellis, K.J., Shypailo, R.J., Pratt, J.A., & Pond, W.G. (1994). Accuracy of dual-energy x-ray absorptiometry for body-composition measurements in children. *The American Journal of Clinical Nutrition*, 60(5), 660-665.

Eyler, A. A., Baker, E., Cromer, L., King, A. C., Brownson, R. C., & Donatelle, R. J. (1998). Physical activity and minority women: a qualitative study. *Health Education & Behavior*, 25(5), 640-652.

Evenson, K. R., Goto, M. M., & Furberg, R. D. (2015). Systematic review of the validity and reliability of consumer-wearable activity trackers. *International Journal of Behavioral Nutrition and Physical Activity*, 12. Retrieved from <http://search.proquest.com.ezproxylocal.library.nova.edu/docview/1779890789?accountid=6579>

- Fields, D.A., Goran, M.I., & McCrory, M.A. (2002). Body-composition assessment via air displacement plethysmography in adults and children: A review. *American Journal of Clinical Nutrition*, 75, 453-467.
- Finkelstein, E.A., Ruhm, C.J., & Kosa, K.M. (2005). Economic causes and consequences of obesity. *Annual Review of Public Health*, 26, 239-257.
- Freedson, P.S., Melanson, E., & Sirard, J. (1998). Calibration of the Computer Science and Applications, Inc. accelerometer. *Medicine & Science in Sports & Exercise*, 30(5), 777-781.
- Ferguson, T., Rowlands, A. V., Olds, T., & Maher, C. (2015). The validity of consumer-level activity monitors in healthy adults worn in free-living conditions: a cross-sectional study. *The international journal of behavioral nutrition and physical activity*, 12. doi: 10.1186/s12966-015-0201-9
- Gao, Z. (2008). Perceived competence and enjoyment in predicting students' physical activity and cardiorespiratory fitness. *Perceptual and Motor Skills*. doi: 10.2466/pms.107.2.365-372
- Garnier, D., & Benefice, E. (2006). Reliable method to estimate characteristics of sleep and physical inactivity in free-living conditions using accelerometry. *Annals of Epidemiology*, 16(5), 364-369.
- Ginde, S.R., Geliebter, A., Rubiano, F., Silva, A.M., Wang, J., Heshka, S., & Heymsfield, S.B. (2005). Air displacement plethysmography: Validation in overweight and obese subjects. *Obesity Research*, 13(7), 1232-1237. doi:10.1038/oby.2005.146

- Going, S. B., Lohman, T. G., Cussler, E. C., Williams, D. P., Morrison, J. A., & Horn, P. S. (2011). Percent body fat and chronic disease risk factors in U.S. children and youth. *American Journal of Preventive Medicine*, 41(4 Suppl. 2), 77-86.
- Goran, M.I. (1994). Applications of the doubly labeled water technique for studying total energy expenditure in young children: A review. *Pediatric Exercise Science*, 6, 11-30.
- Goran, M. I., Poehlman, E. T., & Danforth, E., Jr. (1994). Experimental reliability of the doubly labeled water technique. *American Journal of Physiology*, 29(3), 510-515.
- Gutin, B., Litaker, M., Islam, S., Manos, T., Smith, C., & Treiber, F. (1996). Body-composition measurement in 9-11-y-old children by dual-energy X-ray absorptiometry, skinfold-thickness measurements, and bioimpedance analysis. *American Journal of Clinical Nutrition*, 63(3), 287-292.
- Gutin, B., Yin, Z., Humphries, M.C., & Barbeau, P. (2005). Relations of moderate and vigorous physical activity to fitness and fatness in adolescents. *The American Journal of Clinical Nutrition*, 81, 746-750.
- Hallal, P.C., Victora, C.G., Azevedo, M.R., & Wells, J.C. (2006). Adolescent physical activity and health: A systematic review. *Sports Medicine*, 36, 1019-1030.
- Hands, B., Parker, H., & Larkin, D. (2006). Physical activity measurement methods for young children: A comparative study. *Measurement in Physical Education and Exercise Science*, 10(3), 203-214.
- Hannon, J.C., Ratliffe, T., & Williams, D.P. (2006). Agreement in body fat estimates between hand-held bioelectrical impedance analyzer and skinfold thicknesses in

African American and Caucasian adolescents. *Research Quarterly for Exercise and Sport*, 77(4), 519-526.

Harsha, D.W., & Bray, G.A. (1996). Body composition and childhood obesity.

Endocrinology and Metabolism Clinics of North America, 25(4), 871-885.

Hedley, A.A., Ogden, C.L., Johnson, C.L., Carroll, M.D., Curtin, L.R., & Flegal, K.M.

(2004). Prevalence of overweight and obesity among U.S. children, adolescents, and adults, 1999-2002. *Journal of the American Medical Association*, 291(23), 2847-2850.

Huck, S. W. (2008). *Reading statistics and research* (5th ed.). New York: Allyn and Bacon.

Jago, R., Watson, K., Baranowski, T., Zakeri, I., Yoo, S., Baranowski, J., & Conry, K.

(2006). Pedometer reliability, validity and daily activity targets among 10- to 15-year-old boys. *Journal of Sports Sciences*, 24(3), 241-251.

doi:10.1080/02640410500141661

Kazdin, A.E. (2005). Evidence-based assessment for children and adolescents: issues in measurement development and clinical application. *Journal of Clinical Child and Adolescent Psychology*, 34(3), 548-558.

Keller, B.A. (2008). State of the art reviews: Development of fitness in children: The influence of gender and physical activity. *American Journal of Lifestyle Medicine* 2(1), 58-74. Doi: 10.1177/1559827607308802.

Kirk, S., Zeller, M., Claytor, R., Santangelo, M., Khoury, P.R., & Daniels, S.R. (2005).

The relationship of health outcomes to improvement in BMI in children and adolescents. *Obesity Research*, 13(5), 876-882. doi: 10.1038/oby.2005.102

- Kohl, H., Fulton, J.E., & Capersen, C.J. (2000). Assessment of physical activity among children and adolescents: A review and synthesis. *Preventive Medicine, 31*, S54-S76.
- Kvaavik, E., Tell, G.S., & Klepp, K. (2003). Predictors and tracking of body mass index from adolescence into adulthood. *Archives of Pediatrics and Adolescent Medicine, 157*, 1212-1218.
- Laurson, K. R., Eisenmann, J. C., & Welk, G. J. (2011). Body fat percentile curves for U.S. children and adolescents. *American Journal of Preventative Medicine, 41*(4S2), S87-S92
- Lizzer, S., Bedogni, G., Agosti, F., De Col, A., Mornati, D., & Sartorio, A. (2008). Comparison of dual-energy X-ray absorptiometry, air displacement plethysmography and bioelectrical impedance analysis for the assessment of body composition in severely obese Caucasian children and adolescents. *British Journal of Nutrition, 100*(4), 918-924.
- Léger, L.A., Mercier, D., Gadoury, C., & Lambert, J. (1988). The multistage 20 metre shuttle run test for aerobic fitness. *Journal of Sports Sciences, 6*(2), 93-101.
- Lohman, T.G. (1986). Applicability of body composition techniques and constants for children and youth. *Exercise and Sport Science Reviews, 14*, 325-357.
- Lohman, T. G., Ring, K., Pfeiffer, K., Camhi, S., Arredondo, E., Pratt, C., . . . Webber, L. S. (2008). Relationships among fitness, body composition, and physical activity. *Medicine and Science in Sports and Exercise, 40*(6), 1163-1170. doi: 10.1249/MSS.0b013e318165c86b

- Mash, E.J., & Hunsley, J. (2005). Evidence-based assessment of child and adolescent disorders: Issues and challenges. *Journal of Clinical Child and Adolescent Psychology, 34*, 362-379.
- Martinez-Gomez, D., Ruiz, J. R., Ortega, F. B., Casajus, J. A., Veiga, O. L., Widhalm, K., ... Sjostrom, M. (2010). Recommended levels and intensities of physical activity to avoid low-cardiorespiratory fitness in European adolescents: The HELENA study. *American Journal of Human Biology, 22*(6), 750-6. doi: 10.1002/ajhb.21076
- McClain, J.J., Abraham, T.L., Brusseau, T.A., & 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.
- McKenzie, T.L., Sallis, J.F., & Nader, P.R. (1991). SOFIT: System for Observing Fitness Instruction Time. *Journal of Teaching in Physical Education, 11*, 195-205.
- McNamee, J., & van der Mars, H. (2005). Accuracy of momentary time sampling: A comparison of varying interval lengths using SOFIT [System for Observing Fitness Instruction Time]. *Journal of Teaching in Physical Education, 24*(3), 282-292.
- Mooney, A., Kelsey, L., Fellingham, G.W., George, J.D., Hager, R.L., Myrer, J.W., & Vehrs, P.R. (2011). Assessing body composition of children and adolescents using dual-energy x-ray absorptiometry, skinfolds, and electrical impedance. *Measurement in Physical Education and Exercise Science, 15*, 2-17. doi: 10.1080/1091367X.2011.537873

- Mota, J., Valente, M., Aires, L., Silva, P., Santos, M. P., & Ribeiro, J. C. (2007). Accelerometer cut-points and youth physical activity prevalence. *European Physical Education Review, 13*(3), 287-299. doi: 10.1177/1356336X07081795
- Nilsson, A., Brage, S., Riddoch, C., Anderssen, S. A., Sardinha, L. B., Wedderkopp, N., . . . Ekelund, U. (2008). Comparison of equations for predicting energy expenditure from accelerometer counts in children. *Scandinavian Journal of Medicine & Science in Sports, 18*(5), 643-650.
- Nilsson, A., Ekelund, U., Yngve, A., & Sjoestroem, M. (2002). Assessing physical activity among children with accelerometers using different time sampling intervals and placements. *Pediatric Exercise Science, 14*(1), 87-96.
- Ogden, C. L., Carroll, M. D., Curtin, L. R., Lamb, M. M., & Flegal, K. M. (2010). Prevalence of high body mass index in US children and adolescents, 2007-2008. *Journal of the American Medical Association, 303*(3), 242-249. doi: 1001/jama.2009.2012
- Ogden, C. L., Li, Y., Freedman, D. S., Borruea, L. G., & Flegal, K. M. (2011). *Smoothed percentage body fat percentiles for U.S. children and adolescents, 1999-2004*. Hyattsville, MD: National Center for Health Statistics.
- Ogden, C. L., Yanovski, S. Z., Carroll, M. D., & Flegal, K. M. (2007). The epidemiology of obesity. *Gastroenterology, 132*(6), 2087-2102. doi: 10.1053/j.gastro.2007.03.052
- Ojiambo, R., Cuthill, R., Budd, H., Konstabel, K., Casajús, J. A., González-Agüero, A., . . . Pitsiladis, Y. P. (2011). Impact of methodological decisions on accelerometer

outcome variables in young children. *International Journal of Obesity*, 35(Suppl. 1), 98-103. doi:10.1038/ijo.2011.40

Ortega, F.B., Ruiz, J.R., Castillo, M.J., & Sjöström, M. (2008). Physical fitness in childhood and adolescence: a powerful marker of health. *International Journal of Obesity*, 32, 1-11.

Ortega, F.B., Tresaco, B., Ruiz, J.R., Moreno, L.A., Martin-Matillas, M., Mesa, J.L.,... Castillo, M.J. (2007). Cardiorespiratory fitness and sedentary activities are associated with adiposity in adolescents. *Obesity*, 15(6), 1589-1599.

Pate, R.R., Stevens, J., Pratt, C., Sallis, J.F., Schmitz, K.H., Webber, L.S.,... Young, D.R. (2006). Objectively measured physical activity in sixth-grade girls. *Archives of Pediatrics & Adolescent Medicine*, 160(12), 1262-1268.

Patrick, K., Norman, G. J., Calfas, K. J., Sallis, J. F., Zabinski, M. F., Rupp, J., & Cella, J. (2004). Diet, physical activity, and sedentary behaviors as risk factors for overweight in adolescence. *Archives of Pediatrics & Adolescent Medicine*, 158(4), 385-390.

Plowman, S.A., & Meredith, M.D. (2013). *FitnessGram/ActivityGram Reference Guide*. Dallas, TX: The Cooper Institute.

Pfeiffer, K. A., Mciver, K. L., Dowda, M., Almeida, M., & Pate, R. R. (2006). Validation and calibration of the actual accelerometer in preschool children. *Medicine & Science in Sports & Exercise*, 38(1), 152-157.

doi:10.1249/01.mss.0000183219.44127.e7

- Pope, R.P., Coleman, K.J., Gonzalez, F.B., Barron, F., & Heath, E.M. (2002). Validity of a revised system for observing fitness instruction time (SOFIT). *Pediatric Exercise Science, 14*, 135-146.
- Prentice, A. M., & Jebb, S. A. (2001). Beyond body mass index. *Obesity Reviews: An Official Journal of the International Association for the Study of Obesity, 2*(3), 141-147.
- Putnum, J. J., Allschouse, J. E., & Kantor, L. S. (2002) U.S. per capita food supply trends: more calories, refined carbohydrates, and fats. *Food Review, 25*(3), 2-15.
- Reilly, J. J., Wilson, J., & Durnin, J. V. (1995). Determination of body composition from skinfold thickness: A validation study. *Archives of Disease in Childhood, 73*, 305-310.
- Rennie, K. L., Livingstone, M. B. E., Wells, J. C. K., McGloin, A., Coward, W. A., Prentice, A. M., & Jebb, S. A. (2005). Association of physical activity with body-composition indexes in children aged 6-8 y at varied risk of obesity. *American Journal of Clinical Nutrition., 82*(1), 13-20.
- Ruiz, J. R., Rizzo, N. S., Hurtig-Wennlof, A., Ortega, F. B., Warnberg, J., & Sjöström, M. (2006). Relations of total physical activity and intensity to fitness and fatness in children; the European Youth Heart Study, *American Journal of Clinical Nutrition, 84*, 298-302.
- Sacks, F. M., Bray, G. A., & Carey, V. J. (2009). Comparison of weight-loss diets with different compositions of fat, protein, and carbohydrates. *The New England Journal of Medicine, 360*, 859-873.

Schaefer, S. E., Salazar, M., Bruhn, C., Saviano, D., Boushey, C., & Van Loan, M. D.

(2009). Influence of race, acculturation, and socioeconomic status on tendency toward overweight in asian-american and mexican-american early adolescent females. *Journal of Immigrant and Minority Health, 11*(3), 188-197.

doi:<http://dx.doi.org.ezproxylocal.library.nova.edu/10.1007/s10903-008-9150-6>

Schoeller, D. A. (1988). Measurement of energy expenditure in free-living humans by using doubly labeled water. *Journal of Nutrition., 118*(11), 1278-1289.

Scruggs, P.W., Beveridge, S.K., & Clocksin, B.D. (2005). Tri-axial accelerometry and heart rate telemetry: relation and agreement with behavioral observation in elementary physical education. *Measurement in Physical Education and Exercise Science, 9*(4), 203-218.

Sechrest, L., Stickle, T., & Stewart, M. (1998). The role of assessment in clinical psychology. In M. Hersen & A Bellak (Series Eds.) & C. R. Reynolds (Vol. Ed.), *Comprehensive clinical psychology: Vol. 4. Assessment* (1-32). Tarrytown, NY: Elsevier Science.

Serdula, M. K., Ivery, D., Coates, R. J., Freedman, D. S., Williamson, D. F., & Byers T.

(1993). Do obese children become obese adults? A review of the literature.

Preventative Medicine, 22(2), 167-177. doi: 10.1006/pmed.1993.1014.

Sherry, B., Jefferds, M.E., & Grummer-Strawn, L.M. (2007). Accuracy of adolescent self-report of height and weight in assessing overweight status: A literature review. *Archives of Pediatric & Adolescent Medicine, 161*(12), 1154-1161.

Shephard, R.J., Allen, C., Benade, A.J., Davies, C.T., Di Prampero, P.E., Hedman, R.,...

Simmons, R. (1968). The maximum oxygen intake: An international reference

standard of cardio-respiratory fitness. *Bull World Health Organization*, 38(5), 757-764.

Sirard, J. R., Forsyth, A., Oakes, J. M., & Schmitz, K. H. (2011). Accelerometer test-retest reliability by data processing algorithms: Results from the twin cities walking study. *Journal of Physical Activity and Health*, 8(5), 668-674.

Slaughter, M.H., Lohman, T.G., Baileau, R.A., Horswill, C.A., Stillman, R.J., Van Loan, M.D., & Bamen, D.A. (1988). Skinfold equations for estimation of body fatness in children and youth. *Human Biology*, 60, 709-723.

Steinberger, J., Jacobs Jr, D.R., Ratz, S., Moran, A., Hong, C.P., & Sinaiko, A.R. (2005). Comparison of body fatness measurements by BMI and skinfolds vs dual energy X-ray absorptiometry and their relation to cardiovascular risk factors in adolescents. *International Journal of Obesity*, 29, 1346-1352.

Sun, D.X., Shmidt, G., & Teo-Koh, S.M. (2008). Validation of the RT3 accelerometer for measuring physical activity of children in simulated free-living conditions. *Pediatric Exercise Science*, 20(2), 181-197.

Taber, D. R., Pratt, C., Charneco, E. Y., Dowda, M., Phillips, J. A., & Going, S. B. (2014). Participation in vigorous sports, not moderate sports, is positively associated with cardiorespiratory fitness among adolescent girls. *Journal of Physical Activity and Health*, 11(3), 596-603.

Telford, R.D. (2007). Low physical activity and obesity: Causes of chronic disease or simply predictors. *Medicine & Science in Sports & Exercise*, 39(8), 1233-1240.
doi: 10.1249/mss.0b013e31806215b7

- Thaw, J. M., Villa, M., Reitman, D., DeLucia, C., Gonzalez, V., & Hanson, K. L. (2014). Evidence-based fitness promotion in an afterschool setting: implementation fidelity and its policy implications. *New Directions for Youth Development, 143*, 103-131. doi: 10.1002/yd.20106
- Troiano, R.P. (2006). Translating accelerometer counts into energy expenditure: Advancing the quest. *Journal of Applied Physiology, 100* (4), 1107-1108. doi:10.1152/jappphysiol.01577
- Troiano R.P., Berrigan D., Dodd K.W., Masse L.C., Tilert T., & McDowell M. (2008) Physical activity in the United States measured by accelerometer. *Medicine and Science in Sports and Exercise. 40*(1).181–188.
- Trost, S.G. (2001). Objective measurement of physical activity in youth: Current issues, future directions. *Exercise and Sport Sciences Reviews, 29*, 32-36.
- Trost, S.G. (2007). State of the art reviews: Measurement of physical activity in children and adolescents. *American Journal of Lifestyle Medicine, 1* (4), 299-314.
- Trost, S.G., McIver, K.L., & Pate, R.R. (2005). Conducting accelerometer-based activity assessments in field-based research. *Medicine and Science in Sports and Exercise, 37*(11 Suppl.) 531-543.
- Tudor-Locke, C., Williams, J.E., Reis, J.P., & Pluto, D. (2002). Utility of pedometers for assessing physical activity. *Sports Medicine, 32*(12), 795-808.
- Turner-McGrievy, G. M., Beets, M. W., Moore, J. B., Kaczynski, A. T., Barr-Anderson, D. J., & Tate, D. F. (2013). Comparison of traditional versus mobile app self-monitoring of physical activity and dietary intake among overweight adults

participating in an mHealth weight loss program. *Journal of the American Medical Informatics Association*, 20(3), 513-518.

Tyrrell, V.J., Richards, G., Hofman, P., Gillies, G.F., Robinson, E., & Cutfield, W.S. (2001). Foot-to-foot bioelectrical impedance analysis: A valuable tool for the measurement of body composition in children. *International Journal of Obesity*, 25, 273-278.

Wabitsch, M. (2000a). Overweight and obesity in European children: Definition and diagnostic procedures, risk factors and consequences for later health outcome. *European Journal of Pediatrics*, 159, S8-S13.

Wabitsch, M. (2000b). Overweight and obesity in European children and adolescents: Causes and consequences, treatment and prevention. *European Journal of Pediatrics*, 159(Suppl. 1),5-7.

Ward, D.S., Evenson, K.R., Vaughn, A., Brown Rodgers, A., & Troiano, R.P. (2005). Accelerometer use in physical activity: Best practices and research recommendations. *Medicine & Science in Sports & Exercise*, 1(Suppl), S582-S588.

Welk, G.J., Corbin, C.B., & Dale, D. (2000). Measurement issues in the assessment of physical activity in children. *Research Quarterly for Exercise and Sport*, 71(2 Suppl.), 59-73.

Wheeler, S. & Twist, C. (2010). Methods of assessing body fatness among children: Implications for the National Child Measurement Programme. *European Physical Education Review*, 16(1), 81-93.