RELATION BETWEEN THE FITNESSGRAM® FITNESS ASSESSMENT AND
SELF-REPORTED PHYSICAL ACTIVITY QUESTIONS

Jacob Tucker, B.A.

Thesis Prepared for the Degree of

MASTER OF SCIENCE

UNIVERSITY OF NORTH TEXAS

August 2011

APPROVED:

Scott Martin, Major Professor
James R. Morrow, Jr., Committee Member
Allen Jackson, Committee Member and Chair
    of the Department of Kinesiology,
    Health Promotion and Recreation
Jerry R. Thomas, Dean of the College of
    Education
James D. Meernik, Acting Dean of the
    Toulouse Graduate School
Tucker, Jacob. *Relation between the FITNESSGRAM® fitness assessment and self-reported physical activity questions.* Master of Science (Kinesiology), August 2011, 41 pp., 6 tables, references, 41 titles.

The FITNESSGRAM® is regularly used to assess physical fitness (PF) of adolescents. In addition to the PF assessment, the FITNESSGRAM also includes self-report physical activity (PA) items. The purpose of this study was to examine whether the self-report aerobic, muscular strengthening, and flexibility PA behavior items indicated adolescents’ cardiorespiratory, muscular strength, and flexibility fitness and their body composition. Logistic regression analysis was used to examine the relation between the amount of PA and PF status. Adolescents not meeting the recommended PA amount had significantly higher odds of not achieving a healthy fitness status. Meeting the recommended PA amount was associated with achieving healthy PF status. Thus, adolescents’ amounts of aerobic, muscular strengthening, and flexibility PA were an indication of their corresponding health-related PF standard.
Copyright 2011

by

Jacob Tucker
ACKNOWLEDGEMENTS

This research was supported in part by a grant from the National Association for Sport and Physical Education (NASPE). The author would like to thank the Denton Independent School District middle school teachers and students for their involvement in the project. I would also like to extend appreciation to Dr. Scott Martin, my major professor, for his knowledge and support throughout my thesis work as well as Dr. Allen Jackson and Dr. James R. Morrow, Jr. for their valuable assistance.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>iii</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>vi</td>
</tr>
<tr>
<td>RELATION BETWEEN THE FITNESSGRAM® FITNESS ASSESSMENT AND SELF-REPORTED PHYSICAL ACTIVITY QUESTIONS</td>
<td>1</td>
</tr>
<tr>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>Methodology</td>
<td>5</td>
</tr>
<tr>
<td>Participants</td>
<td>5</td>
</tr>
<tr>
<td>Demographic and Background Information</td>
<td>5</td>
</tr>
<tr>
<td>Instruments</td>
<td>5</td>
</tr>
<tr>
<td>Self-Report Items</td>
<td>5</td>
</tr>
<tr>
<td>Physical Fitness</td>
<td>6</td>
</tr>
<tr>
<td>PACER</td>
<td>6</td>
</tr>
<tr>
<td>Curl-up</td>
<td>7</td>
</tr>
<tr>
<td>90° Push-up</td>
<td>7</td>
</tr>
<tr>
<td>Trunk Lift</td>
<td>8</td>
</tr>
<tr>
<td>Back-Saver Sit and Reach</td>
<td>8</td>
</tr>
<tr>
<td>Shoulder Stretch</td>
<td>9</td>
</tr>
<tr>
<td>Body Mass Index</td>
<td>10</td>
</tr>
<tr>
<td>Statistical Analysis</td>
<td>10</td>
</tr>
<tr>
<td>Results</td>
<td>11</td>
</tr>
<tr>
<td>Correlations</td>
<td>11</td>
</tr>
<tr>
<td>Cardiorespiratory Fitness</td>
<td>12</td>
</tr>
</tbody>
</table>
LIST OF TABLES

<table>
<thead>
<tr>
<th></th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>FITNESSGRAM Physical Activity Questions</td>
<td>19</td>
</tr>
<tr>
<td>2</td>
<td>Correlations for Male and Female Participants</td>
<td>20</td>
</tr>
<tr>
<td>3</td>
<td>Odds Ratios from Logistic Regression Analysis: Cardiorespiratory Fitness</td>
<td>21</td>
</tr>
<tr>
<td>4</td>
<td>Odds Ratios from Logistic Regression Analysis: Muscular Strength Fitness</td>
<td>22</td>
</tr>
<tr>
<td>5</td>
<td>Odds Ratios from Logistic Regression Analysis: Flexibility Fitness</td>
<td>23</td>
</tr>
<tr>
<td>6</td>
<td>Odds Ratios from Logistic Regression Analysis: Body Composition</td>
<td>24</td>
</tr>
</tbody>
</table>
RELATION BETWEEN THE FITNESSGRAM® FITNESS ASSESSMENT AND SELF-REPORTED PHYSICAL ACTIVITY QUESTIONS

Introduction

Adolescents’ physical fitness (PF) is an important health topic. Adolescents’ PF can protect them from diseases associated with sedentary living (Cooper Institute, 2007). Although PF influences health and should be regularly examined, it is often not assessed due to the training and knowledge required, time necessary to complete the fitness testing, and the equipment needed (Harris & Cale, 2007; Silverman, Keating, & Phillips, 2008). Instead of actual fitness testing, some use self-reported physical activity (PA) due to the ease of administration (Sirard & Pate, 2001). The purpose of this study was to analyze the relationship between middle school students’ responses to the FITNESSGRAM® PA questions and whether they meet the corresponding FITNESSGRAM Healthy Fitness Zone™ (HFZ) standard. The aim was to determine whether middle school students’ self-reported amounts of endurance, strength, and flexibility accurately reflected their cardiorespiratory fitness, strength, and flexibility as measured by the FITNESSGRAM. Adolescents should have an accurate self-referenced perspective of the type of fitness behaviors leading to quality physical health. A goal of the FITNESSGRAM program is to promote enjoyable regular PA that leads to good health, function, and well-being throughout a person’s lifetime (Cooper Institute, 2007). Body composition, gender, ethnicity, economic status, and age were analyzed to determine whether these factors are associated with the FITNESSGRAM measures.

PF testing is a necessary aspect of school-based physical education programs that should affect the manner in which physical educators teach health and promote PA (Payne & Morrow, 2009). Although the majority of physical educators utilize some sort of PF testing, it is merely an
isolated part of the physical education program and its reasons for use do not necessarily match those recommended by researchers and test designers (Keating & Silverman, 2004). For PF testing to have a positive influence on school-based health and physical education it should be used regularly as an integral part of physical education and fitness instruction. Students’ knowledge and understanding of PF should increase and they should have a greater potential to meet the basic PF standards (Silverman, Keating, & Phillips, 2008). Once PF testing has been completed there are constructive ways to use the results. Mahar and Rowe (2008) list promoting PA by providing personalized results as a priority following the completion of PF testing. The authors also note that using criterion-referenced standards gives the students PF goals and attainable health standards. Weirisma and Sherman (2008) note that when preparing students for PF testing, physical education teachers should have students regularly participate in appropriate physical education practices, provide opportunities for additional PA during and immediately after school, and conduct proper instruction, training, and practice for each PF test. One of the most important aspects of school-based PF testing is using the assessment as a tool for teaching and promoting enjoyable and regular PA (Weirsma & Sherman, 2008).

Self-report data collection is one of the most common forms of population-level PA measurement (Katzmarzyk & Tremblay, 2007). Previous researchers have noted, however, that self-reports typically demonstrate moderate to low validity when used in child and adolescent populations (Levy & Readdy, 2009; Ridley, Olds, & Hill, 2006). This limitation could be due to recall bias and cognitive ability in youth (Eisenmann, 2003). Discrepancies have been found between self-report PA (SRPA) and actual PA data as measured by accelerometers resulting in poor validity for the self-report questionnaire (Chinapaw, Slootmaker, Schuit, van Zuidam, & van Mechelen, 2009; Slootmaker, Schuit, Chinapaw, Seidell, & van Mechelen, 2009). When
compared to pedometer data, SRPA on walking does not accurately match step count (Scott, Eves, French, & Hoppe, 2007). Parents of adolescents often give more accurate responses about their children’s PA than the children themselves (Le Coq, Boeke, Bezemer, Colland, & van Eijk, 2000; Sithole, & Veugelers, 2008). In all of the previously mentioned studies, adolescents over-report their PA amount. Although there is evidence that SRPA has the potential to be an accurate assessor of actual PA, there are also potential problems with using SRPA data to measure PA in children and adolescents.

Various SRPA surveys have been used with adolescents which have resulted in moderate to high test-retest reliability (ICC = 0.57-0.92; Hume, Ball, & Salmon, 2006; Liu, Wang, Tynjala, Lv, Zhang, & Kannas, 2010). When compared to heart-rate monitor data, one SRPA survey showed strong validity ($r_s = 0.87$) in assessing PA (Belton & Mac Donncha, 2010). Results of studies comparing SRPA to accelerometer data indicate positive correlations (ICC = 0.77, $r = 0.40-0.568$) between the adolescents’ responses and their subsequent PA (Han & Dinger, 2009; Prochaska, Sallis, & Long, 2001; Ridley, Olds, & Hill, 2006; Stanley, Boshoff, & Dollman, 2007). SRPA has also been positively correlated with time spent in PA ($r = 0.51$) and negatively correlated with time spent sedentary ($r = -.45$) as assessed by an activity monitor (Ekelund, Neovius, Linne, & Rossner, 2006). Based on the aforementioned literature, a number of SRPA surveys have shown acceptable reliability and validity. When compared to objective measures of PA, there have been even stronger implications for the use of SRPA data in children and adolescents. The literature reveals contradictory results concerning the use of SRPA data in the assessment of actual PA among children and adolescents. In all of the previously noted studies, the authors examined PA as an outcome variable. However, none addressed whether the participants were healthy based on criterion-referenced standards of PF. Because having quality
physical health is such an important goal concerning children and adolescents, this is a necessary measurement issue. To determine the worth of a SRPA assessment, a measure of PF that objectively classifies whether participants are physically healthy is needed. Recent legislation in the state of Texas provided an opportunity to measure PF and obtain SRPA data.

In 2007 the Texas State House of Representatives and Senate passed legislation requiring each school district to assess the health-related PF of students in grades three through twelve each year (Morrow, Martin, Welk, Zhu, & Meredith, 2010). That same year the Texas Education Agency (TEA) selected the FITNESSGRAM as the assessment tool to accomplish this goal (Cooper et al., 2010). The FITNESSGRAM assessment tool was developed in 1982 by The Cooper Institute in Dallas, Texas (Cooper Institute, 2007). It is a comprehensive PF assessment battery designed specifically for youth. The assessment consists of items to measure aerobic capacity, muscular strength and endurance, and flexibility as well as body composition. The FITNESSGRAM PF assessment is the evaluation tool used as part of the Physical Best program of the National Association for Sport and Physical Education (NASPE), an association within the American Alliance for Health, Physical Education, Recreation and Dance (AAHPERD, see http://www.aahperd.org/naspe/professionaldevelopment/PB_FAQ.cfm). Physical Best is a K-12 program with resources and training appropriate to every grade level and is linked to national education standards that promote lifelong PA and provides accountability for educators (AAHPERD, 2011). The FITNESSGRAM tests cardiorespiratory endurance, abdominal and upper body strength and endurance, trunk extensor strength and flexibility, hamstring flexibility, and upper arm and shoulder girdle flexibility. Body mass index measures weight relative to height (Cooper Institute, 2007). In addition to the PF assessment, there are three SRPA items, which asks students how many days during the past week they engaged in aerobic, strength, and
flexibility activity (Cooper Institute, 2007). The current study analyzed the relationship between participants’ responses to the SRPA items and their PF level as assessed by the FITNESSGRAM. It was hypothesized that the participants’ self-reported responses to the FITNESSGRAM PA questions would be an indication of their corresponding FITNESSGRAM HFZ standard.

Methodology

Participants

Demographic and background information. The participants in this study were 4621 middle school students in grades six through eight in the Denton Independent School District (DISD) in Denton, Texas. DISD provided information on the participants’ gender, economic status, age, and ethnicity. Economic status is based on the federal guidelines that determine which students qualify for free or reduced lunch. The free or reduced lunch qualifications are based on reported family income. Age was defined as the student’s age at completion of the FITNESSGRAM. For ethnicity, the participants were categorized as being either white or non-white. Of the participants, 50.8% were male and 49.2% were female and 40.5% were economically disadvantaged. The adolescents ranged from 10 to 16 years of age ($M = 12.53$, $SD = 1.018$). The majority of the participants were categorized as white (56%). Not all participants had information on their economic status and ethnicity, nor did they all complete the PF testing and respond to the three SRPA items.

Instruments

Self-report items. The SRPA data stemmed from the FITNESSGRAM PA questions (Cooper Institute, 2007). The FITNESSGRAM PA questions ask participants to recall the past seven days of PA in three different activity areas: aerobic, muscular strength and endurance, and flexibility (see Table 1). These questions were printed on the back of each participant’s score
card. Participants indicated this information by circling the number (0-7) of days they participated in each of the activity areas.

*Physical fitness (PF).* The FITNESSGRAM test battery was used to assess the PF of participants (Cooper Institute, 2007). The FITNESSGRAM assesses aerobic capacity (cardiorespiratory fitness), muscular strength and endurance, flexibility, and body composition. After PF measures are collected, participants’ scores are classified into Healthy Fitness Zones (HFZs) or Needs Improvement Zones (NIZs) based on their gender and age. The HFZs and NIZs are criterion-referenced standards based on age and gender that represent the minimum fitness level at which protection against the diseases of sedentary living is achieved (Cooper Institute, 2007). The specific measures that the FITNESSGRAM used to assess PF in this study were the PACER (Progressive Aerobic Cardiovascular Endurance Run), curl-up, 90° push-up, trunk lift, back-saver sit and reach, shoulder stretch, and body mass index.

*PACER.* The FITNESSGRAM PACER assesses aerobic capacity. The objective is to run as long as possible back and forth across a 20-meter space at a pace that is progressive in intensity, meaning it is easier at the beginning than it is at the end as it gets faster each minute (Cooper Institute, 2007). The PACER is set to music that includes the sound of a beep at certain intervals. The participants’ goal is to run across the 20-meter distance and reach the opposite side by touching the line with their foot before the sound of the beep. When they hear the sound of the beep they are to turn around and run to the opposite side. If participants reach one side prior to the sound of the beep, they wait for the beep before running to the opposite side. The first time participants do not reach the line prior the sound of the beep, they should turn around immediately and run to the opposite side in order to get back on pace. The second time participants do not reach the line prior to the sound of the beep, their PACER is completed.
When participants successfully run from one side to the other before the sound of the beep this is considered a “lap” for scoring purposes. Their PACER score is the total number of successful “laps” they completed.

*Curl-up.* The curl-up assesses abdominal strength and endurance in the FITNESSGRAM. The objective is to complete as many curl-ups as possible up to a maximum of 80 at a specified pace (Cooper Institute, 2007). Participants start by lying on a mat, bending the knees at about a 140° angle, feet flat on the floor, legs slightly apart, arms straight with the palms facing down, and the back of the head touching the mat. The fingers are stretched out with the fingertips touching the top of the measuring strip. The measuring strip is a thin, flat strip of material (cardboard, rubber, etc.) that should be 30 to 35 inches in length and a width of 4.5 inches for the age group used in this study. Participants successfully complete a curl-up by curling up slowly so that the fingertips reach the other side of the measuring strip, heels stay on the floor, elbows are not used to assist the curling up, and the movement is completed at the specified cadence of one curl-up every three seconds. The FITNESSGRAM software packet includes a CD that has the specified cadence telling the participants when to come up and when to go back down. Participants’ curl-up assessment is completed when they reach the 80 repetition maximum, can no longer continue, or make their second error. An error occurs when the heels come off the floor, elbows are used to assist the curling up, or the fingertips do not reach to other side of the measuring strip. Their score is the total number of correctly completed curl-ups.

*90° Push-up.* The 90° push-up assesses upper body strength and endurance in the FITNESSGRAM. The objective is to complete as many 90° push-ups as possible at a specified rhythmic pace (Cooper Institute, 2007). Participants begin in a prone position with the hands placed under the shoulders, fingers stretched out, legs straight and slightly apart, and toes tucked
under the legs. To complete the 90° push-up the participants push up off the floor until the arms are straight while keeping the legs and back in a straight line from head to toes throughout the entire movement. Participants must complete the movement at the specified rhythm of one 90° push-up every three seconds. The FITNESSGRAM software packet includes a CD with the specified rhythmic cadence that tells participants when to push up and when to go back down. This exercise is called the 90° push-up because in the down position the elbows should make a 90° angle. Participants’ 90° push-up assessment is completed when they cannot continue or make two form errors. A form error includes not maintaining the specified rhythmic cadence, not maintaining correct body position with a straight back, not extending the arms fully, and not achieving a 90° angle with the elbows on each repetition. Their score is the total number of correctly completed 90° push-ups.

**Trunk lift.** The trunk lift assesses trunk extensor strength and flexibility and is part of the FITNESSGRAM because of its relationship to low back health, specifically vertebral alignment (Cooper Institute, 2007). The objective is to lift the upper body off the floor using the back muscles and to hold the position. Participants begin by lying face down on a mat with the toes pointed and hands placed under the thighs. It is recommended that some kind of marker be placed on the mat in line with their eyes. During the movement participants should focus on the marker so that an accurate measurement can be attained. To complete the trunk lift, participants lift their upper body off the floor in a slow and controlled manner. Their head should be aligned with their spine. Using a ruler the test administrator then measures the distance from the floor to the chin in inches, with a maximum distance of 12 inches.

**Back-saver sit and reach.** The back-saver sit and reach is one of the flexibility assessment items in the FITNESSGRAM. It specifically assesses hamstring flexibility. The objective is to be
able to reach a specified distance on both the right and left sides of body (Cooper Institute, 2007). In this study the Figure Finder Flex-Tester® developed by Novel Products, Inc. (www.novelproducts.com) was the testing apparatus. Participants begin by removing their shoes and sitting down at the testing apparatus. One leg is fully extended with the foot flat against the face of the apparatus while the opposite knee is bent with the sole of the foot flat on the floor. To complete the movement the arms are extended with the palms down, one hand on top of the other, while the participants reach forward four times. The fourth reach should be held for at least one second while the test administrator measures the number of inches reached. After completing the movement on one side, participants switch legs and complete the same steps on the opposite side. To score the back-saver sit and reach the number of inches on each side is recorded to the nearest 1/2 inch. The HFZ is based on the mean between the right and left sides.

Shoulder stretch. The shoulder stretch is another flexibility assessment item in the FITNESSGRAM. It assesses upper arm and shoulder girdle flexibility. The objective is to be able to touch the fingertips on each hand together behind the back by reaching over the shoulder and under the elbow on both the right and left sides (Cooper Institute, 2007). The shoulder stretch is done for both the left and right arms. To complete the shoulder stretch on the right hand side participants reach with the right hand over the right shoulder and down the back. Simultaneously, they reach up with the left hand behind the back trying to touch the fingers of the right hand. Participants then switch hands and do the same movement this time reaching with the left hand down over the left shoulder and the right hand up behind the back trying to touch the fingers of each hand together. Scoring for the shoulder stretch is based on a yes/no criterion. If participants are able to touch the fingertips together on the right hand side a “yes” is recorded. The same goes for the left hand side.
**Body mass index.** Participants’ height and weight were measured to establish body mass index or BMI which is calculated using the following formula:

\[
\text{BMI} = \left( \frac{\text{weight (kg)}}{\text{height}^2 \text{ (m)}} \right)
\]

BMI is an indication of the appropriateness of participants’ weight relative to height (Cooper Institute, 2007). The data were entered in feet and inches for height and pounds for weight then converted and calculated using the metric formula. Participants removed their shoes when measuring height and weight. Fractions of an inch or pound were dropped to the nearest whole number.

**Statistical Analysis**

Correlations between responses to the FITNESSGRAM PA questions and PF status as measured by the FITNESSGRAM were examined. Logistic regression analysis was used to examine the relations of aerobic, strengthening, and flexibility PA with PF status (achieving HFZ or NIZ). Because the 2008 Physical Activity Guidelines (U.S. Department of Health and Human Services, 2008) recommend at least three days per week of aerobic, muscular strength, and bone strength activities for children and adolescents, the data were divided as such. Because the FITNESSGRAM PA questions and fitness assessment examine flexibility instead of bone strength, flexibility fitness was used in the regression model for this study. In the correlation and regression models, participants self-reporting three days or more of aerobic, muscular strengthening, and flexibility PA were separated from those self-reporting less than three days of aerobic, muscular strengthening, and flexibility PA. Therefore, there were two PA groups for each of the three FITNESSGRAM PA questions: those reporting three or more days of the activity and those reporting less than three days of the activity. Achieving healthy cardiorespiratory fitness was defined as being in the HFZ on the PACER. Because there are three
assessment items of muscular strength (90° push-up, curl-up, and trunk lift), achieving healthy muscular strength fitness was defined as being in the HFZ on any two of the three measures. Achieving healthy flexibility fitness was defined as being in the HFZ on both flexibility items (back-saver sit and reach and shoulder stretch). Based on this division, it was anticipated that those in the three or more day’s group would be significantly more likely to achieve a healthy PF level (i.e., be in the HFZ) according to the FITNESSGRAM criteria. Gender, age, ethnicity (white, non-white), and economic status (economically disadvantaged, not economically disadvantaged) were controlled for in the logistic regression model.

Results

Correlations

Pearson correlations were examined to determine the relations between the participants’ responses to the FITNESSGRAM PA questions and PF as assessed by the FITNESSGRAM. Ethnicity, economic status, and age variables were also examined. The data were separated by gender. For both male and female participants, significant positive correlations were found between amounts of aerobic PA and amounts of muscular strengthening PA ($r = .305, p = .001; r = .321, p = .001$; see Table 2). For both male and female participants, significant positive correlations were found between amounts of aerobic PA and amounts of flexibility PA as well ($r = .353, p = .001; r = .367, p = .001$; see Table 2). Additionally, for both male and female participants, significant positive correlations were found between amounts of muscular strengthening PA and amounts of flexibility PA ($r = .371, p = .001; r = .326, p = .001$; see Table 2). For both male and female participants significant positive correlations were found between being in the cardiorespiratory HFZ and being in the BMI HFZ ($r = .347, p = .001, r = .334; p = .001$; see Table 2). Interestingly, for male participants negative correlations were found between
amount of flexibility PA and being in the muscular strength HFZ ($r = -.216, p = .001$; see Table 2). For both male and female participants, significant positive correlations were found between being in the cardiorespiratory HFZ and being in the muscular strength HFZ ($r = .304, p = .001$; $r = .344, p = .001$; see Table 2).

**Cardiorespiratory Fitness**

The FITNESSGRAM PACER was used to assess cardiorespiratory fitness. Healthy cardiorespiratory fitness was defined as being in the HFZ on the PACER. Odds ratios from logistical regression analysis were used to assess the relationship between self-reported amounts of aerobic, muscular strengthening, and flexibility PA and cardiorespiratory fitness. From a seven day recall, those participants self-reporting less than three days of at least 60 minutes of aerobic PA had significantly higher odds of being in the NIZ for cardiorespiratory fitness (OR = 1.910, 95 CI = 1.636 – 2.230, $p = .001$; see Table 3). Those self-reporting less than three days of muscular strengthening PA had even higher odds of being in the NIZ for cardiorespiratory fitness (OR = 3.413, 95 CI = 2.759 – 4.220, $p = .001$; see Table 3). Those self-reporting less than three days of flexibility PA also had significantly higher odds of being in the NIZ for cardiorespiratory fitness (OR = 2.421, 95 CI = 2.072 – 2.829, $p = .001$; see Table 3).

**Muscular Strength Fitness**

The FITNESSGRAM 90° push-up, curl-up, and trunk lift were used to measure muscular strength fitness. Healthy muscular strength fitness was defined as being in the HFZ on any two of these three measures. Odds ratios from logistic regression analysis were used to examine the relationship between self-reported amounts of aerobic, muscular strengthening and flexibility PA and muscular strength fitness. From a seven day recall, those participants self-reporting less than three days of at least 60 minutes of aerobic PA had significantly elevated odds of being in the
NIZ for at least two of the three muscular strength measures (OR = 1.904, .95 CI = 1.586 – 2.286, \( p = .001 \); see Table 4). Those self-reporting less than three days of muscular strengthening PA also had significantly greater odds of being in the NIZ for at least two of the three measures (OR = 3.392, .95 CI = 2.743 – 4.195; \( p = .001 \); see Table 4). Those self-reporting less than three days of flexibility PA had significantly higher odds of being in the NIZ for at least two of the three strength measures as well (OR = 2.812, .95 CI = 2.343 – 3.375, \( p = .001 \); see Table 4).

**Flexibility Fitness**

The FITNESSGRAM back-saver sit and reach and shoulder stretch were used to evaluate flexibility fitness. Healthy flexibility fitness was defined as being in the HFZ for both of these measures. Odds ratios from logistic regression analysis were used to gauge the relationship between self-reported amounts of aerobic, muscular strengthening, and flexibility PA and flexibility fitness. From a seven day recall, those participants self-reporting less than three days of at least 60 minutes of aerobic activity had significantly higher odds of being in the NIZ for at least one of the flexibility measures (OR = 1.703, .95 CI = 1.462 – 1.982, \( p = .001 \); see Table 5). Those self-reporting less than three days of muscular strengthening PA had significantly greater odds of being in the NIZ for at least one of the measures (OR = 1.454, .95 CI = 1.193 – 1.771, \( p = .001 \); see Table 5). Those self-reporting less than three days of flexibility PA also had significantly higher odds of being in the NIZ for at least one of the measures (OR = 1.636, .95 CI = 1.408 – 1.901, \( p = .001 \); see Table 5).

**Body Composition**

Body composition was measured by obtaining participants’ weight and height and then calculating BMI from these measurements. Healthy body composition was defined as being in the HFZ according to the FITNESSGRAM standards for BMI. Odds ratios from logistic
regression analysis were used to assess the relationship between self-reported amounts of aerobic, muscular strengthening, and flexibility PA and body composition. From a seven day recall, those participants self-reporting less than three days of at least 60 minutes of aerobic PA were at significantly higher odds of being in the NIZ for body composition (OR = 1.237, .95 CI = 1.050 – 1.456, p = .011; see Table 6). Those self-reporting less than three days of muscular strengthening PA also had significantly higher odds of being in the NIZ for body composition (OR = 1.319, .95 CI = 1.069 – 1.627, p = .010; see Table 6). Those self-reporting less than three days of flexibility PA had significantly elevated odds of being in the NIZ for body composition as well (OR = 1.284, .95 CI = 1.092 – 1.509, p = .002; see Table 6).

Discussion

PF testing is an important aspect of adolescent health-related PF. The amount of PA that adolescents participate in greatly influences PF test results, thereby serving as an indicator of their health-related PF. It is important for physical educators to use PF testing in a manner that promotes regular PA so that students can achieve basic health standards (Silverman, Keating, & Phillips, 2008). The current study examined the relations between self-reported amounts of PA and PF test results in adolescents. According to correlation data, amounts of aerobic, muscular strengthening, and flexibility PA were found to be positively related. Meeting a healthy cardiorespiratory fitness standard and having a healthy BMI were also found to be directly related. Consistent with the findings of Morrow and Freedson (1994), there was a moderate positive relationship between amount of aerobic PA and meeting a healthy cardiorespiratory fitness standard. In addition, achieving healthy cardiorespiratory fitness and healthy muscular strengthening fitness was positively related.
Odds ratios from logistical regression analysis revealed significant findings. The regression model was set up to demonstrate the odds of not meeting a healthy fitness standard for the FITNESSGRAM assessment items when self-reporting less than three days of PA. However, the inverse can also be explained by the current findings. Participants that reported three or more days in the past week of at least 60 minutes of aerobic PA had significantly higher odds of achieving a healthy cardiorespiratory fitness standard, muscular strength fitness standard, flexibility fitness standard, and a healthy body composition. Participants that reported three or more days in the past week of muscular strengthening PA had even greater odds of meeting a healthy cardiorespiratory fitness standard, muscular strength fitness standard, flexibility fitness standard, and a healthy body composition. Participants that reported three or more days in the past week of flexibility PA had significantly higher odds of achieving a healthy cardiorespiratory fitness standard, muscular strength fitness standard, flexibility fitness standard, and a healthy body composition as well.

Muscular strengthening PA revealed the strongest indication of meeting health-related PF standards. Engaging in at least three days of muscular strengthening PA during the past week resulted in greater odds of achieving healthy PF standards than participating in either aerobic or flexibility PA during the same time period. This is in agreement with Pollock and Vincent (1996) who found that strength training plays a key role in improving many aspects of health (as cited in Haas, Feigenbaum, & Franklin, 2001). One possible explanation for this finding in the current study is that it may be easier for adolescents to recall muscular strengthening PA than aerobic and flexibility PA.

The results of the current study revealed that engaging in three or more days of aerobic (for at least 60 minutes), muscular strengthening, or flexibility PA per week puts participants at
much greater odds of being in the HFZ for any of the FITNESSGRAM assessment measures. Therefore, the hypothesis that the participants’ self-reported responses to the FITNESSGRAM PA questions would be an indication of their corresponding FITNESSGRAM HFZ standard was supported.

*Strengths*

A strength of the FITNESSGRAM as a PF assessment is that it is based on criterion referenced standards. Participants’ scores on each of the exercises are based on their age and gender and converted into them being in the HFZ or the NIZ. It is a comprehensive test battery that includes testing aerobic capacity, muscular strength and endurance, and flexibility. Because there are only three self-reported FITNESSGRAM PA questions that examine the past seven day period it takes minimal time to complete. The FITNESSGRAM offers a test administration manual so that those doing the testing can obtain accurate measurements.

*Limitations*

One limitation to the FITNESSGRAM testing procedure is the reliability of those administering the test. Not all physical education teachers are well trained in administering the FITNESSGRAM assessment. Because of this, measurements are prone to error, causing problems with internal validity. To reduce this potential problem from occurring, a certified FITNESSGRAM administrator was on site overseeing the testing process at all times. The FITNESSGRAM certification process consists of completing an online training course based on the FITNESSGRAM manual. The course implements thorough treatment of the test protocols and takes one through the philosophy of the FITNESSGRAM program. At the end of the course there is an exam. If passed, a certificate of successful completion is presented. The measurement data are likely to be more accurate when a certified FITNESSGRAM administrator is present or
actually administering the test (Mahar & Rowe, 2008). Another potential limitation to the current study is that participants could over-report their PA due to social desirability. However, because the FITNESSGRAM PA questions were on the reverse side of their PF results card, this may have curtailed the possibility of socially desirable responses. By doing the PF testing first, the participants’ SRPA responses may have been more accurate.

Implications

There are a few implications in the current study. The first is that based on the participants’ FITNESSGRAM test results, the self-report FITNESSGRAM PA questions provide an accurate portrayal of adolescents’ amounts of PA. It is important for SRPA data to be validated so that meaningful inferences can be drawn from such data. This provides physical educators with information on the type of PA behaviors that can lead to quality physical health. Another implication of the current study is the amount of PA adolescents should be engaging in to meet basic health-related PF standards. This reveals meaningful associations between PA and PF in adolescents. A healthy level of PF can be achieved by doing more PA. The results of the current study indicate that participating three or more days per week of at least 60 minutes of aerobic PA, three or more days per week of muscular strengthening PA, and three or more days per week of flexibility PA puts adolescents at higher odds of being in the cardiorespiratory, muscular strength, flexibility, and body composition HFZs. Meeting these minimum standards elevates the odds of having a PF level that protects against the diseases of sedentary living. This has important health implications for adolescents. Physical educators should communicate these implications and use PF testing appropriately to assess whether their students are engaging in enough PA to meet basic health-related PF standards. The goal is for adolescents to participate in sufficient PA to obtain higher quality physical health which results in improved quality of life.
Future research could involve an intervention for participants who report less than three days of PA for any of the PA questions and are in the NIZ for any of the FITNESSGRAM assessment items. After a baseline of the FITNESSGRAM test battery is established, participants could be given opportunities to increase their PA to meet the minimum requirement of three days per week for each of the FITNESSGRAM PA question categories. After a sufficient time, these students could be re-tested on the FITNESSGRAM test battery to see if there are improvements in PF scores and a greater likelihood of achieving the HFZ. This within group comparison could demonstrate the influence an intervention that increases the amount of PA has on health-related PF.
Table 1

*FITNESSGRAM Physical Activity Questions*

For each of the following questions, think about what you have done during the past 7 days.

---

1. On how many days were you physically active for a total of at least 60 minutes? This includes moderate activities (walking, slow bicycling, or outdoor play) as well as vigorous activities (jogging, active games, or active sports such as basketball, tennis, or soccer). (Add up all the time you spend in any kind of physical activity that increases your heart rate and makes you breathe hard some of the time.)

   0 Days   1 Day   2 Days   3 Days   4 Days   5 Days   6 Days   7 Days

---

2. On how many days did you do exercises to strengthen or tone the muscles such as push-ups, sit-ups, or weight lifting?

   0 Days   1 Day   2 Days   3 Days   4 Days   5 Days   6 Days   7 Days

---

3. On how many days did you do stretching exercises to loosen up or relax the muscles? This includes exercises such as toe touches, knee bending, and leg stretching.

   0 Days   1 Day   2 Days   3 Days   4 Days   5 Days   6 Days   7 Days
Table 2

Correlations for Male and Female Participants

<table>
<thead>
<tr>
<th></th>
<th>APA</th>
<th>SPA</th>
<th>FPA</th>
<th>CRF</th>
<th>MSF</th>
<th>FF</th>
<th>BMI</th>
<th>ETH</th>
<th>ES</th>
<th>AGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>APA</td>
<td>1</td>
<td>.305**</td>
<td>.353**</td>
<td>.154**</td>
<td>.124**</td>
<td>.090**</td>
<td>.046*</td>
<td>-.013</td>
<td>.017</td>
<td>-.038</td>
</tr>
<tr>
<td>SPA</td>
<td>.321**</td>
<td>1</td>
<td>.371**</td>
<td>.187**</td>
<td>.208**</td>
<td>.071**</td>
<td>.033</td>
<td>-.062**</td>
<td>.074**</td>
<td>-.026</td>
</tr>
<tr>
<td>FPA</td>
<td>.367**</td>
<td>.326**</td>
<td>1</td>
<td>-.216**</td>
<td>.120**</td>
<td>.071**</td>
<td>.012</td>
<td>-.048*</td>
<td>-.098**</td>
<td></td>
</tr>
<tr>
<td>CRF</td>
<td>.099**</td>
<td>.218**</td>
<td>.103**</td>
<td>1</td>
<td>.304**</td>
<td>.159**</td>
<td>.347**</td>
<td>.022</td>
<td>.028</td>
<td>.025</td>
</tr>
<tr>
<td>MSF</td>
<td>.107**</td>
<td>.214**</td>
<td>.149**</td>
<td>.344**</td>
<td>1</td>
<td>.207**</td>
<td>.182**</td>
<td>.010</td>
<td>-.006</td>
<td>.056**</td>
</tr>
<tr>
<td>FF</td>
<td>.113**</td>
<td>.078**</td>
<td>.099**</td>
<td>.182**</td>
<td>.209**</td>
<td>1</td>
<td>.157**</td>
<td>-.036</td>
<td>.063**</td>
<td>-.025</td>
</tr>
<tr>
<td>BMI</td>
<td>.041</td>
<td>.073**</td>
<td>.042</td>
<td>.334**</td>
<td>.208**</td>
<td>.165**</td>
<td>1</td>
<td>-.140**</td>
<td>.103**</td>
<td>-.084**</td>
</tr>
<tr>
<td>ETH</td>
<td>-.130**</td>
<td>-.095**</td>
<td>-.045**</td>
<td>-.082**</td>
<td>-.156**</td>
<td>-.073**</td>
<td>-.144**</td>
<td>1</td>
<td>-.496**</td>
<td>-.018</td>
</tr>
<tr>
<td>ES</td>
<td>.092**</td>
<td>.100**</td>
<td>.080**</td>
<td>.177**</td>
<td>.177**</td>
<td>.118**</td>
<td>.123**</td>
<td>-.476**</td>
<td>1</td>
<td>.060</td>
</tr>
<tr>
<td>AGE</td>
<td>-.059**</td>
<td>.010</td>
<td>-.081**</td>
<td>.270**</td>
<td>.081**</td>
<td>.003</td>
<td>.055**</td>
<td>-.018</td>
<td>.095**</td>
<td>1</td>
</tr>
</tbody>
</table>

Note. Upper diagonal represents males, lower diagonal represents females. *Correlation is significant at the 0.05 level (2-tailed) and **Correlation is significant at the 0.01 level (2 tailed). APA: Aerobic PA; SPA: Strength PA; FPA: Flexibility PA; CRF: Cardiorespiratory fitness; MSF: Muscular strength fitness; FF: Flexibility fitness; BMI: Body mass index; Eth: Ethnicity; ES: Economic status.
Table 3

*Odds Ratios from Logistic Regression Analysis: Cardiorespiratory Fitness*

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Odds Ratio</th>
<th>95% C.I. for OR Lower</th>
<th>95% C.I. for OR Upper</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>PAG Aerobic</td>
<td>3303</td>
<td>1.910</td>
<td>1.636</td>
<td>2.230</td>
<td>.001</td>
</tr>
<tr>
<td>PAG Strength</td>
<td>3312</td>
<td>3.413</td>
<td>2.759</td>
<td>4.220</td>
<td>.001</td>
</tr>
<tr>
<td>PAG Flexibility</td>
<td>3309</td>
<td>2.421</td>
<td>2.072</td>
<td>2.829</td>
<td>.001</td>
</tr>
</tbody>
</table>

*Note.* Controlling for gender, age, ethnicity, and economic status.
Table 4

Odds Ratios from Logistic Regression Analysis: Muscular Strength Fitness

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Odds Ratio</th>
<th>95% C.I. for OR Lower</th>
<th>95% C.I. for OR Upper</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>PAG Aerobic</td>
<td>3260</td>
<td>1.904</td>
<td>1.586</td>
<td>2.286</td>
<td>.001</td>
</tr>
<tr>
<td>PAG Strength</td>
<td>3269</td>
<td>3.392</td>
<td>2.743</td>
<td>4.195</td>
<td>.001</td>
</tr>
<tr>
<td>PAG Flexibility</td>
<td>3267</td>
<td>2.812</td>
<td>2.343</td>
<td>3.375</td>
<td>.001</td>
</tr>
</tbody>
</table>

*Note.* Controlling for gender, age, ethnicity, and economic status.
Table 5

*Odds Ratios from Logistic Regression Analysis: Flexibility Fitness*

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Odds Ratio</th>
<th>95% C.I. for OR Lower</th>
<th>95% C.I. for OR Upper</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>PAG Aerobic</td>
<td>3316</td>
<td>1.703</td>
<td>1.462</td>
<td>1.982</td>
<td>.001</td>
</tr>
<tr>
<td>PAG Strength</td>
<td>3323</td>
<td>1.454</td>
<td>1.193</td>
<td>1.771</td>
<td>.001</td>
</tr>
<tr>
<td>PAG Flexibility</td>
<td>3321</td>
<td>1.636</td>
<td>1.408</td>
<td>1.901</td>
<td>.001</td>
</tr>
</tbody>
</table>

*Note.* Controlling for gender, age, ethnicity, and economic status.
Table 6

*Odds Ratios from Logistic Regression Analysis: Body Composition*

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Odds Ratio</th>
<th>95% C.I. for OR Lower</th>
<th>95% C.I. for OR Upper</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>PAG Aerobic</td>
<td>3342</td>
<td>1.237</td>
<td>1.050</td>
<td>1.456</td>
<td>.011</td>
</tr>
<tr>
<td>PAG Strength</td>
<td>3351</td>
<td>1.319</td>
<td>1.069</td>
<td>1.627</td>
<td>.010</td>
</tr>
<tr>
<td>PAG Flexibility</td>
<td>3349</td>
<td>1.284</td>
<td>1.092</td>
<td>1.509</td>
<td>.002</td>
</tr>
</tbody>
</table>

*Note.* Controlling for gender, age, ethnicity, and economic status.
APPENDIX

REVIEW OF LITERATURE
Fitness Testing in Adolescents

Testing physical fitness (PF) is an integral part of school-based physical education classes and is used extensively by physical educators. With a vast array of PF tests and varying standards, PF assessment is a daunting task. According to the 2006 SHPPS (School Health Policies and Programs Study), 13.7% of states in the U.S. required and 35.3% recommended PF testing for middle school students. Among all states, 35.3% required or recommended using the FITNESSGRAM, 25.5% required or recommended the President’s Challenge Physical Fitness Test, 17.6% recommended the President’s Challenge Health Fitness Test, and 3.9% recommended the YMCA Youth Fitness Test for middle school students (Lee, Burgeson, Fulton, & Spain, 2007). Keating and Silverman (2004) surveyed 325 physical education teachers in 13 different states to investigate the use of PF tests in school-based physical education programs. Results indicated that 83% of the teachers surveyed employ PF tests. Of these, 61% used nationally available tests. The most commonly used test was the President’s Challenge (40.3%), a norm-referenced PF test. Criterion-referenced tests such as the FITNESSGRAM and the Young Men’s Christian Association (YMCA) Youth Fitness Test were utilized by 19.3% and 0.9% of the teachers, respectively. The AAHPERD Physical Best Test was used by 6.6% of teachers and the Chrysler Fund-AAU Physical Fitness Test was used by 2.6%. The most common purpose of testing PF was to teach students to find their own areas of need. The most frequent use of students’ results was to give students a report of their performance. The authors of this study found a discord between researchers’/test designers’ recommendations for use of PF tests and physical education teachers’ use of tests and that PF testing was just an isolated part of school-based physical education programs that did not affect teaching methods (Keating & Silverman, 2004).
Silverman, Keating, and Phillips (2008) offer guidelines for how PF testing can have a positive impact on physical education in schools. First, PF testing should be used as a vital part of PF instruction. The authors note that when it is used as an educational tool, PF testing can promote regular PA and other health-related behaviors. Second, the results yielded by PF testing should have an impact on PF instruction and student learning of PA. Physical educators can teach students about the process of PA, the health-related PF outcomes, and how the two are complimentary through PF testing. Third, the expectation by physical educators that children and adolescents have the potential to meet basic PF standards should be realized. If school-based PF testing is utilized according to these principles there is the potential for students to cultivate better attitudes, skills, knowledge, and PA patterns (Silverman, Keating, & Phillips, 2008).

For PF testing to be effective and the results to be meaningful, appropriate steps must be taken to ensure reliability. Mahar and Rowe (2008) provide examples of the necessary steps to take prior to PF test administration. First, test administrators and scorers must be trained in the correct procedures. If students are to be tested by their peers, allow them to choose their peer group. Make sure that the tests are suitable for those that are being tested. Provide enough equipment to properly administer the tests, make sure it is in good working order, and be organized when initiating the test. Help the participants prepare for the tests by introducing them as exercises for practice prior to the test and have a demonstrator perform the correct technique (Mahar & Rowe, 2008). In addition, the authors list guidelines for how to use the test results. Promote regular PA by providing personalized feedback and use testing as a tool for teaching PA skills and concepts. Use criterion-referenced standards in order to exhibit reachable health standards. Teach students how to track their own levels of PA and PF, and create a testing environment that is positive and meaningful (Mahar & Rowe, 2008). According to the 2006
SHPPS, 61.2% of states in the U.S. and 62.5% of school districts provided funding for staff development or offered staff development to physical educators for administering or using PF tests. This was an increase from 30.6% and 49.8%, respectively, in the year 2000 (Lee, Burgeson, Fulton, & Spain, 2007).

In school-based PF testing, the physical education teachers have a responsibility to properly prepare students. Physical educators must regularly engage students in appropriate physical education practices and provide opportunities to engage in additional PA during and after school (Wiersma & Sherman, 2008). In addition, instruction, training, and practice of each PF test should be performed prior to the actual PF testing. Once these steps have been taken teachers should engage students in the assessment of their personal PF levels (Wiersma & Sherman, 2008). Most importantly, physical education teachers should communicate that promotion of regular and enjoyable PA is the foundation of PF testing.

Strengths of Self-Report Data Collection

One widely used method of collecting PA data is by self-report survey. Self-report data collection is typically inexpensive, takes less time, and uses minimal equipment compared to other methods involved with PA. It is also noninvasive, making it appropriate for large-scale studies, especially among children and adolescents. Much support exists for the use of self-report measurement that makes it one of the most common forms of PA measurement.

Belton and Mac Donncha (2010) found the Youth PA Self-Report (YPAS) to be a valid method of assessing previous day PA for children. In this study, children wore a heart rate monitor and responded to the YPAS the following day. Spearman’s rho correlation between self-reported activity intensity and heart rate was .87 for responses that followed a weekday and .71 for responses that followed a weekend. Corresponding correlations for activity duration were .54 and .68 for weekday and weekend responses, respectively (Belton & Mac Donncha, 2010).
Concerning a longer recall period, a sample of 11-year old children was asked to respond to questions regarding environmental perceptions. A portion of the questionnaire included items on physical environment perceptions. The child was asked how often he or she participated in certain physical activities at home over the past month. The authors found acceptable test-retest reliability for percent agreement (68%-100%) and moderate to good intraclass correlation coefficients (0.72-0.92) for continuous variables (Hume, Ball, & Salmon, 2006). The authors concluded that these self-report items were acceptable for this age group of children. However, only responses to the survey items were analyzed, not actual PF. In a similar study, the authors used the Health Behavior in School-aged Children (HBSC) self-report questionnaire to examine test-retest reliability. This study was aimed for adolescents at age 11 or age 15. The authors found that the test-retest reliability for the items concerning PA ranged from moderate (ICC = 0.57) to almost perfect (ICC = 0.82) agreement (Liu, Wang, Tynjala, Lv, Zhang, & Kannas, 2010). The recall period for the PA items was one week. Overall, the authors found the test-retest reliability to be satisfactory, but there was no assessment of actual PF.

Positive implications have been found for a SRPA measure that used a recall period of three days and compared responses to actual PA data. Stanley, Boshoff, and Dollman (2007) had a sample of adolescent females (N = 20) wear the Computer Science and Applications, Inc. 7164 accelerometer for seven consecutive days. The participants then completed the 3-day PA Recall (3dPAR) questionnaire (Pate, Ross, Dowda, Trost, & Sirard, 2003) and compared responses with the accelerometer data. Significant correlations were found between self-reported vigorous activity and accelerometer vigorous activity for three days of monitoring (r = 0.49, p < 0.05) and seven days of monitoring (r = 0.57, p < 0.01) (Stanley, Boshoff, & Dollman, 2007). The use of another three day PA recall questionnaire known as the 3DR also revealed strong relationships
between self-reported vigorous activity and accelerometer-derived time when assessed in young adults (Han & Dinger, 2009).

Ekelund, Neovius, Linne, and Rossner (2006) developed a self-report last seven day PA questionnaire for adolescents in order to compare responses to actual PA. The participants wore the MTI activity monitor (Manufacturing Technology Inc., Fort Walton Beach, FL, USA; model WAM 6471) for a minimum of five days. The participants then completed the Swedish Adolescent Physical Activity Questionnaire (SAPAQ) (Ekelund, Neovius, Linne, & Rossner, 2006). SRPA was negatively correlated with time spent sedentary \( (r = -0.45) \) and positively correlated with time spent in PA \( (r = 0.51) \). The authors regarded the SAPAQ as a valid method of assessing PA in adolescents (Ekelund, Neovius, Linne, & Rossner, 2006). Another self-report instrument that has demonstrated high validity and reliability is the Multimedia Activity Recall for Children and Adolescents (MARCA). The MARCA is a previous day recall questionnaire that addresses PA in time slices of five minutes or more. When compared to the MTI Actigraph accelerometer (Model AM7164-2.2C, Manufacturing Technologies, Inc., Fort Walton Beach, FL) data, the MARCA showed high test-retest reliability and comparable validity to other one day self-report questionnaires (Ridley, Olds, & Hill, 2006).

Based on the previously mentioned research, considerable support exists for the use of self-report instrumentation when assessing PA in adolescents. This support has been demonstrated in test-retest and intraclass correlation settings and when compared to actual PA as evaluated by heart rate monitor and accelerometer data. This support has also been displayed across multiple recall time periods. There still remain discrepancies, however, between responses of SRPA and measured PF data that limit its use (Katzmarzyk & Tremblay, 2007).
Weaknesses of Self-Report Data Collection

The effectiveness of self-report data collection is dependent on response accuracy. Despite evidence that the self-report method does not capture accurate levels of PA, it continues to be one of the most common forms of population-level PA measurement (Katzmarzyk & Tremblay, 2007). To accurately assess the value of SRPA time, a comparison to objectively measured time spent doing PA must be conducted. Slootmaker, Schuit, Chinapaw, Seidell, and van Mechelen (2009) compared self-reported time by questionnaire to time measured by an accelerometer spent on PA at moderate and vigorous intensity. Adolescents were given the PAM accelerometer (PAM, model AM101, PAM B.V., Doorwerth, The Netherlands) to wear for 14 consecutive days during waking hours. Following this, the participants completed the Activity Questionnaire for Adults and Adolescents (AQuAA). The adolescents self-reported significantly more time spent on both moderate ($M_{diff} = 596 \text{ min/wk}$) and vigorous ($M_{diff} = 178 \text{ min/wk}$) PA than was reported on the accelerometer (Slootmaker, Schuit, Chinapaw, Seidell, & van Mechelen, 2009). In addition, only 7 to 22% of the self-reported time spent on moderate and vigorous PA was confirmed by the accelerometer. The authors stated that the inaccurate recall and overestimation of time spent doing PA may be because of intermittent daily activities involving breaks or rest periods.

The AQuAA has a past seven day recall period in order to combat recall bias, but has demonstrated poor validity when compared to accelerometer data (Chinapaw, Slootmaker, Schuit, van Zuidam, & van Mechelen, 2009). These authors suggested that because adolescents underestimated sedentary behavior time and overestimated PA time, problems with recall accuracy were still an issue when completing the AQuAA. The International Physical Activity Questionnaire (IPAQ) is a self-report questionnaire that uses a last seven day recall period.
Construct validity of the IPAQ was generally low ($\rho = .02$ to .32), particularly for total PA (Levy & Readdy, 2009). The authors suggested that the accuracy of the IPAQ needs to be improved when assessing PA.

Using the Theory of Planned Behavior (TPB), Scott, Eves, French, and Hoppe (2007) assessed how well a self-report questionnaire on walking predicts actual step count. Participants wore a pedometer (New Lifestyles NL-2000, New Lifestyles Inc., Kansas City, MO) during all waking hours for one week. The participants then completed the seven-day PA recall (Sallis, et al., 1985). The results did not show a significant correlation between self-report measures and pedometer measures of walking. The authors suggested that the participants were largely unaware of their walking behaviors making it difficult to accurately report (Scott, Eves, French, & Hoppe, 2007).

When assessing PA behaviors in children and adolescents by self-report, the question of whether parents provide a more accurate evaluation has been investigated. In one study, children who self-reported more PA and less sedentary activity than what their parents perceived were actually more likely to be overweight or obese (Sithole & Veugelers, 2008). In this study, parent reports had a stronger association between greater PA and healthy weight for children. The authors contended that parent reports provide a more accurate assessment of PA than children’s self-reports (Sithole & Veugelers, 2008). In another parent-child self-report comparison study, le Coq, Boeke, Bezemer, Colland, and van Eijk (2000) gave asthmatic children age’s eight to twelve a quality of life questionnaire that, in part, addressed PA. A parent-report version of the same questionnaire demonstrated better reproducibility and responsiveness. The authors noted that the parents were more sensitive in observing their children’s behavior than the children.
themselves and that in longitudinal studies data should be obtained from parents (le Coq, Boeke, Bezemer, Colland, & van Eijk, 2000).

Another potential method of improving SRPA accuracy among children and adolescents concerns the time format of the questionnaire. Rifas-Shiman et al. (2001) compared an annual format questionnaire to a seasonal format questionnaire. These questionnaires yielded the average number of hours per week spent in total, moderate, and vigorous PA over the previous year. The adolescents in this study over-reported PA on the annual format questionnaire. Among ten year old children, an average of 3.4 more hours per week of PA was reported in the annual format compared to the seasonal format. The authors concluded that due to seasonal variability in PA patterns, a seasonal format questionnaire may improve self-report accuracy when assessing child and adolescent PA (Rifas-Shiman et al., 2001).

Recommendations for Self-Report Data Collection

Based on previously mentioned research, self-report instruments as an assessment of PA among children and adolescents have had mixed results. Major threats to validity of self-report measures include cognitive and situational factors (Brener, Billy, & Grady, 2003). Among cognitive factors, the most prevalent problem involves recall memory. Many times respondents are asked to recall many separate activities as well as the duration and intensity of those activities which can be difficult for adolescents. Situational factors revolve around social desirability bias. Since PA is highly valued, adolescents may over-report PA in an attempt to meet expectations. In one literature review, studies of 61 different PA questionnaires for children and adolescents were analyzed. Results of these studies were evaluated using a standardized checklist known as the Qualitative Attributes and Measurement Properties of Physical Activity Questionnaires (QAPAQ). According to the review, none of the 61 questionnaires showed both acceptable
reliability and validity (Chinapaw, Mokkink, van Poppel, van Mechelen, & Terwee, 2010). However, the authors did find many promising questionnaires whose usefulness could be improved if evaluated in more studies. In a review of the existing literature, Brener, Billy, and Grady (2003) found moderate to substantial reliability for several self-report PA questions as well as moderate to high reliability for several PA measures. Glasgow, Ory, Klesges, Cifuentes, Fernald, and Green (2005) recommend items from the PACE+ project (Prochaska, Sallis, & Long, 2001). The PACE+ project items, which assess PA frequency in the last seven days, have been validated with a sample of middle school students. These items were reliable (ICC = 0.77) and significantly correlated ($r = .40$, $p < .001$) with accelerometer data (Prochaska, Sallis, & Long, 2001). In addition, these items also demonstrated acceptable measurement properties (sensitivity 71%, specificity 63%; Prochaska, Sallis, & Long, 2001).

Other self-report instruments that appear to be appropriate for children and adolescents include the Self-Administered Physical Activity Checklist (SAPAC) and the Physical Activity Questionnaire for Children (PAQ-C) (Bates, 2006). The SAPAC measures previous day PA before, during, and after school. The PAQ-C assesses the frequency of PA that occurs in or outside of school. Chinapaw, Mokkink, van Poppel, van Mechelen, and Terwee (2010) included these two questionnaires as having promise for future studies. Bates (2006) also notes that an objective assessment of PA such as pedometer measures should be used in addition to self-report to get an accurate evaluation of PA. In a systematic literature review, Adamo, Prince, Tricco, Connor-Gorber, and Tremblay (2009) concur with this appraisal. Based on implications from 83 different studies, these authors concluded that because of discrepancies between the two, direct measures of PA should be included with indirect measures in order to provide a more accurate assessment of PA in children and adolescents. Due to increasing health concerns regarding
inadequate PA among adolescents, objective PA monitors are being used more often to address this problem (Esliger & Tremblay, 2007). When coupled with self-report data collection methods, objective measures can maximize PA assessment and evaluate the accuracy of the self-report measure. Based on the existing literature regarding this topic, research studies that employ these guidelines have the potential to provide a more precise assessment of PA in children and adolescents.
REFERENCES


Champaign, IL: Human Kinetics.


