AN ABSTRACT OF THE DISSERTATION OF

So-Yeun Kim for the degree of Doctor of Philosophy in Exercise and Sport Science presented on August 8, 2006.
Title: Sources of Variability in Daily Physical Activity for Secondary Students With and Without Developmental Disabilities

Abstract approved:

Joonkoo Yun

Generalizability theory was used to examine sources of variability in daily physical activity levels of secondary students with and without developmental disabilities (DD), and to determine minimum number of days required for monitoring their typical physical activity levels. Sixteen participants with DD (*M* = 16.7 years, *SD* = 2.7 years), and thirty-one (*M* = 12.3 years, *SD* = 0.5 years) children without DD participated in the study. They wore two pedometers and two accelerometers during 5 weekdays and 4 weekend days. Sources of variability were examined using a two-facet fully crossed design. Twelve separate two-way ANOVAs were employed for each population, physical activity device, and measurement periods (weekday, weekends, and weekdays and weekends combined).

For participants with DD, variance components of the person, and the person by day interaction were the primary sources of variability in daily physical activity for pedometers and accelerometers across weekdays, weekend days, and weekdays and weekends combined. To determine the typical physical activity level with

generalizability coefficients of .80, at least 4, 6, and 8 days of measurement using a pedometer were required during weekdays, weekend days, and weekdays and weekends combined, respectively. Using an accelerometer, at least 4 days of measurements were needed across weekdays, weekends, and weekdays and weekends combined.

For participants without DD, the primary sources of variability during weekdays and weekends were related to variance components of the persons and the person by day interaction for both pedometers and accelerometers. When weekdays and weekends were combined, relatively large percentages of variability were associated with the residual, indicating three way interaction, plus unexplained error. Using one pedometer, to achieve generalizability coefficients of .80 in the measurement of daily physical activity, a minimum number of 5 and 9 days of measurements during weekday and weekends were estimated, respectively. Using one accelerometer, at least 4 days and 14 days of monitoring physical activity were required during weekdays, and weekdays and weekends combined, respectively. However, an estimation of typical physical activity levels during weekdays and weekends combined, using one pedometer as well as during weekends using one accelerometer was unfeasible due to the number of days required for measurement.

Sources of Variability in Daily Physical Activity
for Secondary Students With and Without Developmental Disabilities

by
So-Yeun Kim

A DISSERTATION

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Doctor of Philosophy

Presented August 8, 2006
Commencement June 2007

Philosophy of Doctoral dissertation of So-Yeun Kim presented on August 8, 2006.

APPROVED:

Major Professor, representing Exercise and Sport Science

Chair of Department of Nutrition and Exercise Sciences

Dean of the Graduate School

I understand that my dissertation will become part of the permanent collection of Oregon State University libraries. My signature below authorized release of my dissertation to any reader upon request.

So-Yeun Kim, Author

CONTRIBUTION OF AUTHORS

Dr. Joonkoo Yun was involved with the design and writing of dissertation, and Dr. Jeffrey McCubbin assisted in editing of the final dissertation.

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

LIST OF FIGURES

LIST OF TABLES

LIST OF APPENDICES

SOURCES OF VARIABILITY IN DAILY PHYSICAL ACTIVITY FOR SECONDARY STUDENTS WITH AND WITHOUT DEVELOPMENTAL DISABILITIES

Chapter 1: General Introduction

The emphasis on the importance of regular physical activity is based on the significant relationship between physical inactivity and the risk factors for obesity, coronary heart disease, diabetes, and colon cancer (Blair, 1993; Pate et al., 1995; U.S. Department of Health and Human Service [USDHHS], 1996). The Surgeon General's report on physical activity and health concluded that moderate physical activity can substantially reduce the risk of developing or dying from health related disease such as diabetes, colon cancer, and high blood pressure (USDHHS, 1996). Also, Healthy People 2010 (USDHHS, 2000) identified physical activity as one of the ten leading health indicators.

The importance of regular physical activity for optimal health has led to an increased interest in assessing habitual or typical physical activity behaviors (USDHHS, 1996). To determine typical physical activity patterns, and to evaluate the effectiveness of health promotion programs to increase physical activity, representative data of individuals' typical physical activity should be analyzed. One of the challenges, of assessments to determine typical physical activity, is the high degree of variability in physical activity levels within and among individuals in free living settings.

Identifying the source and magnitude of variability in typical physical activity is important to making of appropriate research designs and interpretation of the results of studies related to physical activity and health. The variability can be divided into biological variability and analytic variability: Biological variability in physical activity occurs when physical activity behavior changes naturally from one measurement period to the next, while analytic variability occurs because instruments provide inconsistent data under the same set of circumstances (Dale, Welk, & Matthews, 2002). Although in the previous literature suggested a term “biological variability”, “behavioral variability” was used in this study. The word, “biological”, refers to “the natural processes of living things” (American Heritage, 2002). However, physical activity is a behavior, and variability in physical activity from one measurement period to the next is a behavioral change rather than a change in the functioning of the body’s process.

Intra-individual variability in physical activity refers to differences in physical activity levels in a person from day to day (Baranowski & de Moor, 2000). Intra-individual variability should be considered in the measurement of habitual physical activity because physical activity levels of an individual can vary from day to day, and such variability can make estimation of habitual physical activity difficult (Baranowski & de Moor, 2000; Levin, Jacobs, Ainsworth, Richardson, & Leon, 1999). Intra-individual variability can be estimated by identifying the number of days of measurement required to get reliable and representative information of usual physical activity in a given population (Gretebeck & Montoye, 1992; Levin et al., 1999).

The minimum number of days required to capture representative data of people’s typical physical activity has been examined using the intra-class correlation (ICC). The minimum number of days was identified using the criterion value for acceptable reliability as ICC $\geq .80$ (Baranowski & de Moor, 2000). Different recommendations have been proposed depending on the population and the instruments that were used for the studies. Based on data collected with pedometers, monitoring 5 days for elementary students (Vincent & Pangrazi, 2002) and 6 days for middle school students (Rowe, Mahar, Raedeke & Lore, 2004) were recommended. Uni-axial accelerometers have also been used to determine the minimum number of days required for monitoring moderate and vigorous physical activity levels (Janz, Witt, & Mahoney, 1995; Trost, Pate, Freedson, Sallis, & Taylor, 2000). Trost et al. (2000) recommended 5 days of monitoring children (Grade 1-6) and 9 days of monitoring adolescents (Grade 7-12), while Janz, Witt, and Mahoney (1995) suggested 6 days of data collection for children between 7 and 15 years old. Similar results were also found for adults. Gretebeck and Montoye (1992) suggested monitoring adults for 5 days using waist pedometers, 6 days using ankle pedometers, and 6 days using uni-axial accelerometers.

Another source of variability that should be considered in the assessments of typical physical activity is inter-instrument variability (i.e., differences between instruments of the same brand), which is analytical variability (Bassett & Strath, 2002). Understanding how much variability exists between units and how variable responses exist over time is important because it affects the accuracy of data (Welk, Schaben, & Morrow, 2004). Inter-instrument variability has been examined using a mechanical setup that allows for a standardized amount of movement, or by comparing outputs from instruments worn on opposite hips of the same individuals.

Acceptable inter-instrument reliability in accelerometers has been found using a mechanical setup (e.g. mechanical shaker table). The ICC of .97 among 9 different

Tritrac tri-axial accelerometers (Kochersberger, McConnell, Kuchibhatla, & Pieper, 1996), a small coefficient of variation, of 1.79% across 4 different Tritrac tri-axial accelerometers (Nichols, Morgan, Sarkin, Sallis & Calfas, 1999), and 5% error among 7 CSA uni-axial accelerometers (Metcalf, Curnow, Evans, Voss, & Wilkin, 2002) were reported. Also, acceptable inter-instrument reliability between two units of accelerometers and pedometers worn on opposite hips has been reported. ICCs ranging from .73 to .87 for Tritrac tri-axial accelerometers (Nicholas et al., 1999) and ICCs of .87 for CSA uni-axial accelerometers (Trost, Ward, Moorehead, Watson, Riner, & Burke, 1998) during treadmill running and walking were examined.

For pedometers, mixed results have been found in different brands of pedometers. The Yamax and Eddie Bauer pedometers showed very close agreements when two pedometers were worn on opposite hips during walks on 4.88 km side walk courses. However, two Pacer pedometers worn on opposite hips recorded significantly different numbers of steps (Bassett et al., 1996).

Only limited information regarding the inter-instrument reliability of pedometers and accelerometers in daily physical activity is available. Barfield, Rowe and Michael (2004) examined inter-instrument reliability of the Yamax Digi-walker pedometers in elementary school-aged children. Each participant in the study wore the Yamax Digi-Walker pedometers on the right and left side of hip for 7 days. High ICCs (> .90) were found for the whole week, during classroom time, during recess time, and during physical education.

To date, intra-individual variability and inter-instrument variability have been examined separately, but both sources of error can occur simultaneously while assessing typical physical activity. Using generalizability theory (G-theory), an estimate of the participants' variability across time as well as systematic errors in days of measurement and measurement errors due to unsystematic sources can be examined (Shavelson & Webb, 1991). In other words, using G-theory the different sources of potential variability (i.e. days and instruments), called facets, can be included in one study. Moreover, the estimate of the relative amount of total variance in the score associated with each facet, and with the interactions among them can be determined from one study (Goodwin, 2001). Not only the major sources of variability, the total magnitude of variability, and a reliability coefficient (Shavelson & Webb, 1991) but also the optimal measurement protocols that minimize the error for a particular condition can be provided by G-theory (Shavelso, Webb, & Rowley, 1989; Ulrich & Wise, 1984). However, the theory remains relatively to be used little in the measurement of physical activity.

Two published studies examined variability of physical activity in treadmill walking (Welk, Schaben & Morrow, 2004) and variability of daily physical activity (Coleman & Epstein, 1998) using the G-theory. In the study by Welk, Schaben and Morrow (2004), two-facets (4 monitors, 3 trials) in a fully crossed design was employed. The largest percentage of variance was associated with differences among participants (from 38.5% to 63.4%). The relatively large amount of variance examined in the trial by person interaction component (from 14% to 21%) and the variance component among monitors, trials and persons (from 18% to 23%). The results of this study suggest that correct and consistent positioning of accelerometers across trials is important in order to decrease measurement error. Coleman and Epstein (1998) examined the variability of daily physical activity levels of sedentary males using accelerometers and self-report. A one-facet, the person by day, design was employed. The largest variance component was the residual term for both accelerometers (46%) and self-reports (56%), indicating unknown sources of error not measured in the study. The results indicate that more complicated designs, containing more than one facet, should be used to examine variability of daily physical activity for sedentary males.

In the assessment of habitual physical activity, the degree of intra-individual variability may differ for individuals with disabilities compared to individuals without disabilities. Sedentary populations may have less intra-individual variability in daily physical activity compared to the regular population. The daily physical activity levels of individuals with peripheral vascular disease were examined over two 2-day periods, separated by 1 week (Sieminki, Cowell, Montgomery, Pillai, & Gardner, 1997). The ICC of .86 for steps per day was found. The result suggests that a shorter sampling period may be sufficient for a sedentary population. Tudor-Locke and Myers (2001) reported that further studies are needed to examine the optimal number of days for monitoring physical activity for sedentary populations although they may need fewer measurement days than active groups. It has been assumed that children with developmental disabilities (DD) have lower daily physical activity levels than children without disabilities although there is only limited information regarding the daily physical activity levels of children with DD (Fernhall, 2002). Significant differences in the daily physical activity levels between children with and without Prader-Willi Syndrome (PWS), aged 4 – 19 years, were examined (Eiholzer et al., 2003). Based on the results of activity protocols performed by their parents, the children with PWS walked less (11.1 km vs. 24.6 km) ($p < .05$), and participated in lower levels of physical activity (256 points vs. 274 points) ($p < .01$) than those without disabilities.

Therefore, the main purpose of this study was to examine the source of variability in measuring typical physical activity levels of secondary students with and without DD using pedometers and accelerometers. The secondary purpose of this study was to examine minimum number of days required for monitoring daily physical activity levels of secondary students with and without DD. Comparisons between the two populations were not made for the following reasons: 1) the age groups of students with and without DD were different, and 2) sample sizes of two populations were different.

Research Questions

The following research questions were examined in this study:

1. How much of the variance in daily physical activity levels in secondary students with and without DD is related to the days?
2. How much of the variance in daily physical activity levels in secondary students with and without DD is related to the instruments?
3. How many days of monitoring physical activity are required to determine habitual or typical physical activity levels of secondary students with and

without DD using pedometers and accelerometers?

Assumptions

1. Participants did not change their physical activity behaviors while wearing pedometers and accelerometers.
2. Participants wore pedometers from the time that they got up in the morning until bedtime.
3. Participants consistently wore pedometers and accelerometers in the correct place.
4. Placement of the two pedometers and the two accelerometers were close enough to capture same body movement.

Delimitations

1. This study was delimitated to secondary students with DD at two school districts in Oregon, and to secondary student without disabilities at a middle school in Oregon.
2. Students with DD were classified by the student's school district according to eligibility under Individuals with Disabilities Education Act.

Limitations

1. The participants were volunteers.
2. Environmental/contextual factors were not measured.

Chapter 2

SOURCES OF VARIABILITY IN DAILY PHYSICAL ACTIVITY FOR SECONDARY STUDENTS WITH DEVELOPMENTAL DISABILITIES

So-Yeun Kim

Abstract

The purpose of this study was to examine sources of variability in the measurement of daily physical activity levels of children with developmental disabilities (DD), and to determine the optimal number of days required for monitoring daily physical activity levels to capture their typical physical activity. Sixteen middle school students with DD (M = 16.7 years, SD = 2.7 years) participated in this study. Participants wore two pedometers and two accelerometers for 9 days including 5 weekdays and 4 weekend days. A two-facet in fully crossed design using six separate two-way ANOVAs was employed to estimate sources of variability in physical activity across weekdays (W), weekends (WK), and weekdays and weekend days combined (WWK) for each device. During W, WK, and WWK, the primary sources of variability were related to the person (53.16%, 39.06%, and 32.95%, respectively) and the person by day interaction components (44.04%, 57.95%, and 52.61%, respectively) using pedometers. Using accelerometers, also large amounts of variability were related to the person (55.29%, 51.06%, and 49.88%, respectively) and the person by day interaction components (25.76%, 38.04%, and 37.13%, respectively) across W, WK, and WWK. Four, six, and eight days of measurements were required to determine typical physical activity levels of children with DD during W, WK, and WWK, respectively. Using one accelerometer, the estimated number of required days of monitoring during W, WK, and WWK was 4 days.

Sources of Variability in Daily Physical Activity for Secondary Students with Developmental Disabilities

Two U.S. government documents, the report of the Surgeon General on physical activity and health (U.S. Department of Health and Human Service [USDDHHS, 1996]) and Health People 2010 (USDHHS, 2000), emphasized health promotion by increasing daily physical activity levels of individuals regardless of gender or age. The emphasis on the importance of regular physical activity is based on significant positive effects of regular physical activity on health and disease (Blair, 1993; Pate et al., 1995, USDHHS, 1996). Healthy People 2010 identified physical activity as one of the ten leading health indicators, and made several objectives related to regular physical activity in children and adults (USDHHS, 2000). Moreover, both U.S. government documents emphasized the importance of health promotion in individuals with disabilities and decreasing health disparities between individuals with and without disabilities, through increasing the daily physical activity levels of individuals with disabilities (USDHHS, 1996; USDHHS, 2000).

It has been reported that sedentary lifestyle of individuals with disabilities contributes to an increased risk of early morbidity and mortality (Pate et al., 1995). Daily physical activity levels of children with developmental disabilities (DD) have been assumed to be lower than children without disabilities although limited research on daily physical activity of children with DD is available (Fernhall, 2002). Daily physical activity levels of children with Prader-Willi Syndrome (PWS) compared to daily physical activity levels of children without disabilities were examined using proxy reports by their parents (Eiholzer et al., 2003). Children with PWS walked less (11.1 km vs. 24.6 km) ($p < .05$), and participated in lower levels of physical activity (256 points vs. 274 points) ($p < .01$) than those without disabilities.

To create effective health promotion programs that increase daily physical activity to decrease the health disparity between children with DD and without disabilities, accurate data on daily physical activity levels of all individuals should be analyzed. The first step is to identify the source and magnitude of variability in daily physical activity levels. It has been reported that the high degree of variability in physical activity within and among individuals in free living settings makes it hard to determine the daily physical activity levels of individuals regardless of age, and gender (Montoye & Taylor, 1984).

Variability of daily physical activity can be due to intra-individual variability and inter-instrument variability. Intra-individual variability in physical activity occurs because physical activity behavior changes naturally from one measurement period to the next (Bassett & Strath, 2002). Intra-individual variability refers to differences in physical activity level in a person from day to day (Baranowski & de Moor, 2000). Intra-individual variability should be considered in the measurement of daily physical activity because physical activity levels of an individual can vary from day to day, and such variability can make the estimations of daily physical activity difficult (Baranowski & de Moor, 2000; Levin, Jacobs, Ainsworth, Richardson, & Leon, 1999). Inter-instrument variability occurs because the same instruments provide inconsistent data under the same set of circumstances (Bassett & Strath, 2002). Understanding how much variability exists between units and how variable responses are over time is important because it affects the accuracy of data (Welk, Schaben, & Morrow, 2004).

To date, very limited information is available regarding to intra-individual variability and inter-instrument variability in daily physical activity levels of children with DD. Several researchers examined intra-individual variability in the daily physical activity of individuals without disabilities by identifying the minimum number of monitoring days to capture representative information of daily physical activity in children using pedometers (Rowe, Mahar, Raedeke & Lore, 2004; Vincent & Pangrazi, 2002) and using uniaxial accelerometers (Janz, Witt, & Mahoney, 1995; Trost, Pate, Freedson, Sallis, & Taylor, 2000) as well as adults using pedometers and uniaxial accelerometers (Gretebeck & Montoye, 1992). However, there are two issues applying the results to children with disabilities.

First, children with DD may have less intra-individual variability in daily physical activity compared to children without disabilities. It has been assumed that children with developmental disabilities (DD) have lower daily physical activity levels than children without disabilities although there is only limited information regarding the daily physical activity levels of children with DD (Fernhall, 2002). Tudor-Locke and Myers (2001) reported that further studies are needed to examine the minimum numbers of days required for monitoring physical activity of sedentary populations because they may need fewer measurement days than active groups. Second, the previous studies did not examine inter-instrument variability, differences between same instruments of the same brand. Inter-instrument variability has been examined using a mechanical setup that allows for a standardized amount of movement, or by comparing outputs from instruments worn on opposite hips of the same individuals. However, intra-individual variability and inter-instrument variability can occur simultaneously while assessing daily physical activity. Both sources of variability should be examined at the same time.

Therefore, the purpose of this study was to examine the sources of variability in daily physical activity levels of children with DD, and to determine minimum numbers of days required for monitoring typical physical activity levels of children with DD.

Method

Participants

A convenience sample of 16 secondary school students between ages of 11 and 20 years with DD (M = 16.7 years, SD = 2.7 years) participated in this study. The participants were recruited from four secondary schools in two school districts, in the rural areas of Northwestern United States. Participants received education in special education classes, and did not have physical disabilities. Special education teachers verified this condition based on the diagnosis of the school districts because the participant's personal information was not accessible to the researcher, and the schools did not provide IQ scores due to school district policies.

Among the participants, there were nine students with mental retardation, two students with Down syndrome, two students with autism, one student with traumatic brain injury, and two students with developmental delay. This study was approved by a

University Institution Review Board prior to recruiting participants. All the participants and their parents/guardians signed written consent forms.

Instruments

Pedometers

Omron HJ-112 pedometers (Omron Healthcare, Vernon Hills, IL) were used to measure number of steps that each participant accumulated throughout the day. This pedometer (2.13 in. x 0.63 in. x 2.88 in., 1.13 oz.) has unique features to count steps according to the manufacturer's claim. First, this pedometer does not count steps until it has registered over four seconds of movement, which can eliminate the chance of counting non step movements as steps. Second, this pedometer is capable of counting steps correctly even when the front of the main unit is placed at an angle of more than 60 degrees from the ground as well as when it is horizontal to the ground. This feature may accurately measure steps of individuals with high waist circumference because this pedometer may be less susceptible to errors that occur due to tilt**.** Moreover, the pedometer can store steps for seven days. It has an internal clock that automatically resets the counts to zero, so users do not have to press the reset button every day.

Pedometers were used after examining the proper calibration using the "shake-test" developed by Vincent and Sidman (2003). Each pedometer was placed in a shipping box from the manufacturer, and shaken 100 times; In order to decrease abnormal movements, the box was moved in a vertical direction while the bottom of the box remained in contact with a table surface. The shake-test was performed twice for each pedometer before and after the study. No pedometers exceeded ± 5% error, ± 5 steps out of 100.

<u>Accelerometers</u>

The accelerometers used to measure the acceleration of body movement were Actiwatches (Mini-Mitter, Bend, OR). This small and light watch-like device (one square inch, .6 oz.) contains an omnidirectional sensor that is sensitive to motion in all directions (Mini-Mitter, 2000). An increase in the degree of speed and motion produces an increase in voltage, and the device stores this information as activity counts. The device can accumulate activity counts based on a sampling time from 15 seconds to 2 minutes. In this study, 30 seconds interval was used.

To ensure the consistency among the accelerometers, prior to data collection, two researchers wore six devices on the waist for 3 hours. Each device was worn at least 10 times. Ten accelerometers, which exceeded ± 15% errors, were recalibrated by the manufacturer. The following equation was used to calculate error;

$$\frac{X_i - [(X_1 + X_2 + X_3 + X_4 + X_5 + X_6)/6]}{X_i} \times 100$$

Note: Where *X* represents the average activity counts of an accelerometer, and *i* represents the sample average.

Procedures

All the participants wore two pedometers, attached to a Velcro belt, on their waists. The pedometers were placed in line with the middle of the right thigh of each participant to examine consistency of the instruments. Also, two accelerometers were both placed on the wrist of the dominant hand of each participant with a non-removable plastic wrist band. Participants were instructed not to cut the band unless they felt discomfort. Data from accelerometers and pedometers were collected by a researcher at least every two days.

Consistent placement of pedometers was emphasized to the participants and their parents. All participants were instructed not to tamper with the instruments in order to prevent inflating counts. Moreover, pedometers were sealed to prevent inflating counts or accidentally resetting the instruments. Participants were asked to wear pedometers and accelerometers for 11 days including 7 weekdays and 4 weekend days between April and June. The first two weekdays were for familiarizing participants with wearing instruments. Each participant was asked to wear two pedometers from the time that they got up in the morning until bedtime, except swimming and showering, but they worn accelerometers throughout the entire data collection. To remind participants to wear the devices every day, the researcher gave two reminder posters (Appendix H) to each participant. Participants were asked to attach the posters wherever they would be seen easily. Also, the researcher gave reminder phone calls every morning to 14 participants who chose to accept phone calls.

To get more detailed information about the participants' physical activity, parents/guardians of each participant were asked to keep a daily physical activity log. They were asked to write the duration, location, and type of physical activity in which the participants engaged. Physical activity logs were collected daily. Weekend logs were collected on Mondays.

Data Analysis

Generalizability theory (G-theory) was used in this study. G-theory allows researchers to estimate reliability while identifying multiple sources of error separately in a single model (Goodwin, 2001; Morrow, 1989; Shavelson, Webb, & Rowley, 1989). It provides not only a single coefficient but also magnitudes of variability due to different source of error (Shavelson & Webb, 1991). The sources of error (e.g., trials, occasions, raters) are called facets, and the facet levels are called conditions (Shavelson, Webb, & Rowley, 1989). Two types of studies can be conducted using G-theory; generalizability study (G-study) and decision study (D-study) (Morrow, 1989). A G-study is designed to provide estimates of the variance components associated with each facet and their interactions (Morrow, 1989; Shavelson, Webb, & Rowley, 1989). A D-study is designed to make substantive decisions about a measurement protocol using the results of the G-study (Morrow, 1989; Shavelson, Webb, & Rowley, 1989).

<u>G-study</u>

Research questions for the source of variability were examined using a two-facet fully crossed design using six separate two way ANOVAs. Two separate 2 by 5 (instrument × day) ANOVAs were employed to examine sources of variability in weekday physical activity levels for pedometers and accelerometers, respectively. For weekend days, a 2 by 4 (instrument × day) ANOVA was employed separately for two different devices. Finally, two separate 2 by 9 (instrument × day) ANOVAs were employed to examine sources of variability in daily physical activity levels (weekdays and weekends combined) for two different devices. Instruments and days were considered random facets. Fully crossed designs indicate that all of the facets are treated as random, and all participants are crossed with all other facets in the model (Morrow, 1989). The facets are treated as random if facets are considered interchangeable with others (Shavelson & Webb, 1991).

Pedometers and accelerometers are not exchangeable because they measure different aspects of physical activity behaviors. Therefore, the data from these two devices were analyzed separately. Weekdays and weekends may not be exchangeable because physical activity patterns during weekdays and weekend days may be different. However, no conclusive information is available on the difference in the physical activity patterns of children with DD between weekdays and weekends. Therefore, the data were analyzed in three different ways; weekdays, weekends, and weekdays and weekend days combined.

Seven sources of variability (Table 2.1) were estimated from the ANOVA results; the seven sources of variability include variance associated with persons, days, and instruments, three two-way interactions (person by day, person by instrument, and day by instrument), and the residual term (three-way interaction plus error). The sources of variability were determined using the VARCOMP procedure from SAS. The percentage of variance associated with each source of variability was calculated by dividing each variance estimate by the total variance.

Table 2.1. Sources of Variability in the Two-Facet Design

Source of Variability	Type of Variability	Variance Notation
Persons (p)	Universe-score variance (object of measurement)	σ^2_p
Instruments (i)	Constant effect for all persons due to stringency of instruments	σ^2_i
Days (d)	Constant effect for all persons due to the inconsistencies of their physical activity behaviors from one day to another	σ^2_d
p × i	Inconsistencies of outputs of instruments on particular persons' physical activity behaviors	σ^2_{pi}
p × d	Inconsistencies from one day to another in particular persons' physical activity behaviors	σ^2_{pd}
i × d	Inconstant effect from all persons due to differences in instruments' stringency from one day to another	σ^2_{id}
p × i × d, *e*	Residual consisting of the unique combination of p, i, d; unmeasured facets that affect the measurement; and/or random events σ	$\sigma^2_{pid,e}$

From the estimated variance components, generalizability coefficients were also calculated. Like reliability coefficient on classical test theory, generalizability coefficients indicates how accurate the generalization is from an individual's observed score is to his or her true universe score (Shavelson & Webb, 1991). Two different coefficients can be calculated in G-theory based on absolute and relative decisions. Absolute decisions are based on the absolute level of performance, and use phi coefficients (Shavelson, Webb, & Rowley, 1989).Relative decisions are based on the relative standing of participants, and use G coefficients (Shavelson, Webb, & Rowley, 1989). The equation for G and phi coefficients were determined as follows (Shavelson

& Webb, 1991);

$$G = \sigma_p^2 / (\sigma_p^2 + \sigma_{pd}^2 / n_d + \sigma_{pi}^2 / n_i + \sigma_{pid,e}^2 / n_i \ n_d)$$

$$phi = \sigma_p^2 / (\sigma_p^2 + \sigma_i^2 / n_i + \sigma_d^2 / n_d + \sigma_{pd}^2 / n_d + \sigma_{pi}^2 / n_i + \sigma_{id}^2 / n_i \ n_d + \sigma_{pid,e}^2 / n_i \ n_d)$$

D-study

Research questions related to minimum number of days required for monitoring daily physical activity were examined using a D-study. By increasing or decreasing the number of facet levels (days and instruments), the minimum number of facet levels required to establish the desired generalizability was determined.

Results

Descriptive statistics for each device across weekdays, weekends, and weekdays and weekend days are presented in Table 2.2. Participants with complete data from pedometers and accelerometers across weekdays, weekends, and all days combined were used. The average steps per weekday, weekend day, and weekday and weekend days were 8299, 5858, and 7106 steps, respectively. The average activity counts per weekday, weekend day, and weekday and weekend days measured with accelerometers were 502744, 437097, and 473567 activity counts, respectively.

Table 2.2. Descriptive statistics for pedometers and accelerometers

	Pedometer (*N* = 16)		Accelerometers (*N* = 15)	
	Mean	SD	Mean	SD
Weekdays	8,299	2,433	502,744	128,456
Weekends	5,858	4,098	437,097	142,804
Weekdays/Weekends	7,106	3,604	473,567	138,670

Pedometers

G-study

Variance component estimates and their relative magnitude for daily physical activity levels measured with pedometers are presented in Table 2.3. During weekdays, the largest source of variability was the person (53.16%), and the second largest source of variability was the person by day interaction (44.04%). During weekends, the largest source of variability was the person by day interaction (57.95%), and the second largest source of variability was the person (39.06%). During weekdays as well as weekends, the other variance components were very low. When data of weekdays and weekend days were combined, the primary sources of variability were the person by day interaction (52.61%) and the person (32.93%) while the day facet was also associated with 13.71% of total variance.

The estimated generalizability coefficients for the average data from two pedometers across five weekdays was high ($G = .85$ $phi = .85$), while moderate generalizability coefficients ($G = .73$, $phi = .72$) was estimated for the average data of two pedometers across four weekend days. When data from weekdays and weekend days were combined, high generalizability coefficients ($G = .85$, $phi = .82$) was also estimated for the average data from two pedometers across 9 days.

D-study

Minimum numbers of days required for monitoring daily physical activity levels was determined with D-study (Table 2.4). To achieve generalizability coefficients of

.80 in the measurement of daily physical activity during weekdays, at least four days of measurements with a pedometer were estimated. During weekends, six days of measurements with two pedometers or seven days of measurements with a pedometer were estimated. When data of weekdays and weekend days were combined, eight days of measurement using a pedometer was estimated. Figure 2.1 presents estimated phi coefficients as the number of measurement day were increased using one pedometer.

Table 2.3. Variance Component Estimates and their Relative Magnitudes for Pedometers

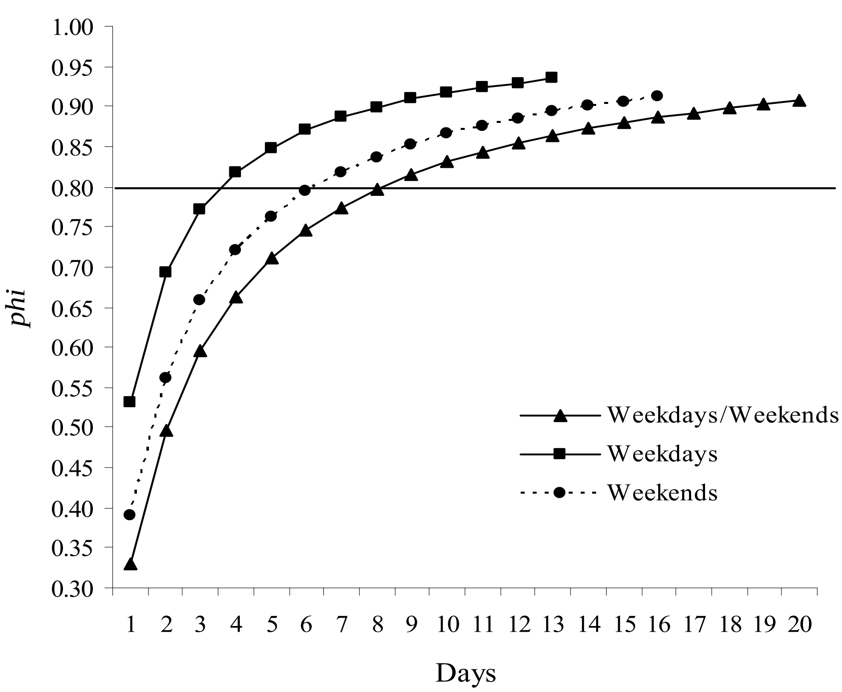
Variation	Estimated Variance Components	Relative Magnitude[1]
Weekdays/Weekends		
Persons (p)	4175921.2	32.93%
Instruments (i)	-888.6[2]	0%
Days (d)	1737938.8	13.71%
p × i	2533.5	0.02%
p × d	6671011.2	52.61%
i × d	910	0.01%
Residual (p × i × d, *e)*	91410.7	0.72%
Weekdays		
Persons (p)	3261406	53.16%
Instruments (i)	-3199.1[2]	0%
Days (d)	-16674.7	0%
p × i	9120.6	0.15%
p × d	2701789.9	44.04%
i × d	3237.8	0.05%
Residual (p × i × d, *e)*	159979.0	2.61%
Weekends		
Persons (p)	6779396	39.06%
Instruments (i)	-570.8[2]	0%
Days (d)	402553.3	2.32%
p × i	-8747.4[2]	0%
p × d	10058437	57.96%
i × d	-1708.3[2]	0%
Residual (p × i × d, *e)*	116528.2	0.67%

Note. [1] Relative magnitude was calculated using estimated variance divided by the total variance. [2] Negative variance components were set to zero in subsequent calculations as suggested by Morrow (1989).

Table 2.4. Estimated Generalizability Coefficients for the Pedometers

Number of Days	Number of Instruments	*G*	*phi*
Weekdays/Weekends			
9	1	.85	.82
8	2	.83	.80
8	1	.83	.80
Weekdays			
5	1	.85	.85
4	1	.82	.82
3	2	.78	.78
Weekends			
7	1	.82	.82
6	2	.80	.80
6	1	.80	.79

Figure 2.1. Estimated phi coefficients using one pedometer



Accelerometers

G-study

Table 2.5 shows estimated variance components and relative magnitude for accelerometers. Similar results found across the days including weekdays, weekend days, and when weekdays and weekend days were combined. During weekdays, the largest source of variability was the person (55.29%) while the person by day interaction and residual were associated with 25.76% and 18.83% of the total variance, respectively. During weekends, relatively large variances were due, to the person (51.06%), to the person by day interaction (25.76%), and to the residual (18.83%). When data of weekends and weekday days were combined, primary sources of variability were due to the person (49.88%) and to the person by day interaction (37.13%) while the residual was associated with 8.85% of the total variance.

High generalizability coefficients were estimated across all three data sets. High generalizability coefficients were estimated for the average of two accelerometer data from five weekdays (G = .89, *phi* = .89) and four weekends days (G = .82, *phi* = .82). Also, when data of weekdays and weekend days were combined, high generalizability coefficients (G = .92, *phi* = .91) were estimated for the average data from two pedometers across nine days.

D-study

Minimum numbers of days for monitoring daily physical activity levels using one or two accelerometers were estimated with the D-study. The results of the D-study are presented in Table 2.6. To achieve generalizability coefficients of .80 in the measurement of daily physical activity during weekdays, at least four days of measurements with an accelerometer or three days of measurement with two accelerometers were estimated. During weekends, four days of measurements with an accelerometer was estimated. Also, when data from weekdays and weekends were combined, four days of measurement using an accelerometer was required. Estimated phi coefficients as the number of measurement day were increased using one accelerometer are presented in Figure 2.2.

Table 2.5. Variance Component Estimates and their Relative Magnitudes for Accelerometers

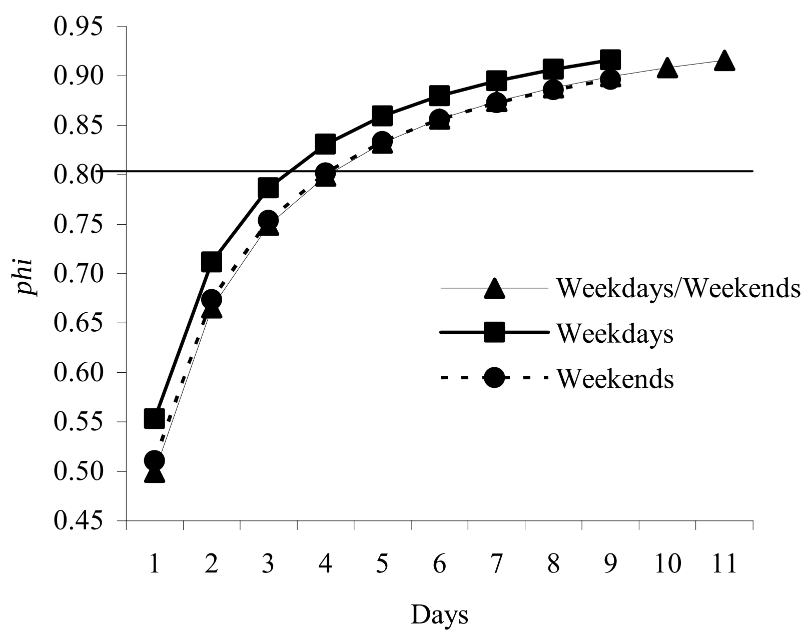
Variation	Estimated Variance Components	Relative Magnitude[1]
Weekdays/Weekends		
Persons (p)	9899151911	49.88%
Instruments (i)	5730049	0.33%
Days (d)	814948547	4.11%
p × i	-110304610[2]	0 %
p × d	7369254010	37.13%
i × d	-88153991[2]	0%
Residual (p × i × d, *e)*	1756067204	8.85%
Weekdays		
Persons (p)	9867018610	55.29%
Instruments (i)	20628176	0.12%
Days (d)	-207514380[2]	0%
p × i	-397096597[2]	0%
p × d	4597618718	25.76%
i × d	-168991272[2]	0%
Residual (p × i × d, *e)*	3359469266	18.83%
Weekends		
Persons (p)	11015271724	51.06%
Instruments (i)	-51433925[2]	0%
Days (d)	-227525937[2]	0%
p × i	113226132	0.53%
p × d	8206200964	38.04%
i × d	38379675	0.18%
Residual (p × i × d, *e)*	2201200641	10.20%

Note. [1] Relative magnitude was calculated using estimated variance divided by the total variance. [2] Negative variance components were set to zero in subsequent calculations as suggested by Morrow (1989).

Table 2.6. Estimated Generalizability Coefficients for Accelerometers

Number of Days	Number of Instruments	*G*	*phi*
Weekdays/Weekends			
5	1	.84	.83
4	2	.83	.81
4	1	.81	.80
Weekdays			
5	1	.86	.86
4	1	.83	.83
3	2	.82	.82
Weekends			
5	1	.83	.83
4	2	.82	.82
4	1	.80	.80

Figure 2.2. Estimated phi coefficients using an accelerometer



Discussion

Pedometers

The results of the G-study indicate that the variance component of the person and the person by day interaction were the primary sources of variability in daily physical activity levels of secondary students with DD when data of weekdays and weekend days were combined. The largest percentage of variation was associated with the person by day interaction component (52.61%), which suggests that some participants accumulated a large number of steps on someday, but not on others. This variation may be due to illness, or participation in physical education classes or after school physical activity on particular days. According to the daily physical activity logs, four participants reported participation in practice for Special Olympics programs on the different days. Some participants attended both adapted physical education classes and regular physical education classes on the particular days. A relatively large percentage of variation was associated with the person (32.93%) and day (13.71%). The variance component of person indicates individual differences among the participants in their daily physical activity levels, while the day facet suggests that something transpired on a particular occasion that affected all participants in the same way, increasing or decreasing their physical activity. This could be due to variations in weather on a given day.

The magnitude of the other estimated sources of variability were negligible. The variance component of instruments was 0% indicating consistent steps between two pedometers. The person by instruments, and instrument by day variance components were very low (0.02%), indicating consistent differences between pedometers across participants and days. The variation associated with the residual was very small (0.72%), which indicates little random measurement error.

Similar results were found when data from weekdays and weekend days were analyzed separately. The largest percentage of the variation was associated with differences in physical activity levels among the participants (53.16%) during weekdays while the variation associated with the person by day was 44.04%. During weekends, the variation associated with the person by day interaction was 57.95%, and the person facet was associated with 39.06% of the total variation. Again, the magnitude of the other estimated sources of variability were negligible.

The large proportion of variance associated with the person and the person by day components as well as very small proportion of variance related to the instruments, the person by instrument, and the instrument by day, indicate two things. First, that major source of variability in the daily physical activity levels of secondary students with DD are related to behavioral variability, indicating physical activity levels vary naturally from day to day. Second, it indicates that very small amount of systematic error was associated with pedometers in the measurement of daily physical activity levels of secondary students with DD in the free living settings exists.

The minimum number of days required for monitoring daily physical activity levels was determined with a D-study based on the results of the G-study. In practice, it is unlikely that two pedometers would be used to measure the daily physical activity levels of individuals, so the only results with one pedometer are discussed. Also, only the estimated phi coefficients based on absolute decisions were discussed because the relative standing of the participants in this study had no bearing on the results; a participant's daily physical activity did not depend on how physically active his or her peers were. The results of D study indicate that at least 8 days of monitoring physical activity levels (*phi* = .80) are needed to determine typical physical activity levels of secondary students with DD using a pedometer, based on the assumption that the daily physical activity pattern during weekdays and weekends are same. Fewer days of measurement are required to determine typical physical activity levels of the participants during weekdays or weekends compared to weekdays and weekends combined. This would be useful information for professionals who may be interested in the promotion of daily physical activity only during weekdays or weekends. To determine daily physical activity levels during weekdays, measurements need to be obtained from at least 4 weekdays (*phi* = .82) while 7 weekend days (*phi* =.82) of measurements are required to determine typical physical activity levels during weekends.

There have been no studies that examined sources of variability in daily physical activity levels of secondary students with DD. However, the study by Rowe et al. (2004) examined variability in daily physical activity levels of middle school students without disabilities. The study estimated that at least six days (R = .83) of monitoring were needed to capture typical physical activity patterns of middle school students without disabilities using intra-class correlation coefficients with data from weekdays and weekends combined. Although a direct comparison should not be made between the studies because the study by Rowe et al. (2004) examined only one source of variability (day), this study found that a greater number of measurement (at least 8 days of measurements when data from weekdays and weekends were combined) than the results of the previous study. It could be explained with difference of mean steps during weekdays and weekends. In this study, the participants accumulated a greater number of average steps during weekdays (8,299) than weekends (5,858) while the previous study found very similar average steps during weekdays (9,504) and weekends (9,005).

Accelerometers

Like the results of G-study for pedometers, large behavioral variability and small analytic variability were also estimated for accelerometers. When data from weekdays and weekend days were combined, the primary sources of variability were the variance component of the person (49.88%) and the person by day interaction (37.13%). Relatively small variances were associated with the day (4.11%) as well as the residual (8.85%). The results indicate that variability in daily physical activity when weekdays and weekends were combined was due to students' natural behavior variability from day to day, rather than inter-instrument variability.

Similar results were found when the accelerometer data were analyzed separately. During weekdays, the largest percentage of variation was associated with differences among the participants (55.29%) while the person by day interaction was related to 25.76% of the total variation. Also, a relatively large percentage of that variation was associated with the residual component (18.83%). During weekends, the primary sources of variability were associated with the person (51.06%) and with the person by day interaction (38.04%). The residual was associated with 10.20% of the total variance. Again, the results suggest that variability in the daily physical activity during weekdays and weekends was due to the changes in behaviors of the participants, rather than measurement errors related to the accelerometers. The residual variance, 18.83% for weekdays and 10.20% for weekends, suggest that variability occurred across days and instruments, and the differences vary by the participants. Also, it could be due to an unknown source of error which was not measured in the study (Shavelson & Webb, 1991). Because the participants wore accelerometer for 24 hours, there might be more sources of variability than "days" and "instruments".

The minimum number of days required to capture typical physical activity levels of secondary students with DD using accelerometers was fewer than using pedometers. The results of the D-study estimated that at least 4 days of monitoring daily physical activity using an accelerometer were required to determine daily physical activity levels (*phi* $\geq$.80) whether data from weekday and weekend days were combined or separated.

No published studies have examined the minimum number of days required to monitor typical physical activity of secondary students with DD using G-theory. However, different recommendations have been proposed to examine moderate and vigorous physical activity levels of children without disabilities using the intra-class correlation. Using uni-axial accelerometers, Trost et al. (2000) recommended 5 days of monitoring children (Grade 1-6) and 9 days of monitoring adolescents (Grade 7-12), whereas Jan, Witt, and Mahoney (1995) suggested 6 days of data collection for children between 7 and 15 years old. Although the results of these studies do not allow comparison with the current results because neither study examined intra-individual variability and inter-instrument reliability of daily physical activity level at the same time. There might be less variability in daily physical activity levels of secondary students with DD compared to those without disabilities, when intensity levels of movement were measured.

Summary

Primary sources of variability in the daily physical activity levels of secondary students with DD were the person and the person by day interaction components. The results indicated that some children had consistently high or low levels of physical activity across days (environment) whereas other children were more sensitive to daily changes. Different profiles of physical behavior patterns of children with DD should be identified. Their profiles may help to design effective intervention strategies to promote daily physical activity levels of children with DD.

Relatively low variance related to inconsistent data between devices, and inconsistent data between device for particular persons and from one day to another day indicate that high reliability evidence of Omron HJ pedometers and Actiwatches for the measurement of physical activity in secondary students with DD during free living settings. Further research should examine validity evidence of both devices in the daily physical activity measurement during free living settings.

Shorter monitoring days were estimated necessary to determine typical physical activity levels for secondary students with DD using accelerometers than pedometers. This result suggests that intensity levels of body movements in children with DD vary less across weekdays, weekends, and weekdays and weekend combined. Examining whether they engage in low intensity levels of physical activity consistently across days, or in enough moderate and vigorous physical activity (MVPA) would be a good addition because of significant relationship between MVPA and health.

References

Baranowski, T., & de Moor, C. (2000). How many days was that? Intra-individual variability and physical activity assessment. *Research Quarterly for Exercise and Sport, 71*(2), 74-78.

Bassett, D. R., & Strath, S. J. (2002). Use of pedometers to assess physical activity. In G. J. Welk (Ed.), *Physical activity assessments for health-related research (pp.163-177)*. Champaign, IL: Human Kinetics.

Blair, S. N. (1993). 1993 C. H. McCloy research lecture: Physical activity, physical activity, and health. *Research Quarterly for Exercise and Sport, 64,* 365-376.

Dale, D. D, Welk, G. J., & Matthews, C. E. (2002). Methods for assessing physical activity and challenges for research. In G. J. Welk (Ed.), *Physical activity assessments for health-related research (pp.19-34)*. Champaign, IL: Human Kinetics.

Eiholzer, U., Nordmann, Y., L'Allemand D., Schlumpf, M., Schmid, S., & Kromeyer-Hauschild, K. (2003). Improving body composition and physical activity in Prader-Willi syndrome. *The Journal of Pediatrics, 142*, 73-78.

Fernhall, B. (2002). Physical activity, metabolic issues, and assessment. *Physical Medicine and Rehabilitation Clinics of North America*, *13*, 925-947.

Goodwin, L. D. (2001). Interrater agreement and reliability. *Measurement in Physical Education and Exercise Science, 5*(1)*,* 13-34.

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

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

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

Minimitter Company Inc. (2000). *Actiwatch activity monitors: Instruction manual*. Sunriver, OR: Author.

Montoye, H. J. & Taylor, H. L. (1984). Measurement of physical activity in population studies: a review. *Human Biology*, *5*, 195-216.

Morrow, J. R. (1989). Generalizability theory. In M.J. Safrit & T.M. Wood (Eds.), *Measurement concepts in physical education and exercise science (pp.73-96).* Champaign, IL: Human Kinetics.

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

Shavelson, R. J. & Webb, N. M. (1991). *Generalisability theory: A primer*. Newbury Park, CA: Sage.

Shavelson, R. J., Webb, N. M., & Rowley, G. L. (1989). Generalizability theory. *American Psychologist, 44*(6), 922-932.

Sieminki, D. J., Cowell, L. L., Montgomery, P. S., Pillai, S. B., & Gardner, A. W. (1997). Physical activity monitoring in patients with peripheral arterial occlusive disease. *Journal of Cardiopulmonary Rehabilitation, 17*, 43-47.

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

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

Tudor-Locke, C. E., & Myers, A. M. (2001). Methodological considerations for researchers and practitioners using pedometers to measure physical (ambulatory) activity. *Research Quarterly for Exercise and Sport, 72*, 1-12.

U.S. Department of Health and Human Services. (1996). *Physical activity and health: A report of the Surgeon General.* Atlanta, GA: U.S. Department of HHS.

U.S. Department of Health and Human Services. (2000). *Healthy people 2010.* Washington, DC: U.S. Gov. Printing Office.

Vincent, S. D., & Sidman, C. L. (2003). Determining measurement error in digital pedometers. *Measurement in Physical Education and Exercise Science, 7*, 19-24.

Vincent, S. D., & Pangrazi, R. P. (2002). Does reactivity exit in children when measuring activity levels with pedometers? *Pediatric Exercise Science, 14,* 56-63.

Welk, G. J., Schaben, J. A., & Morrow, J. R. (2004). Reliability of accelerometry-based activity monitors: A generalizabilty study. *Medicine and Science in Sports and Exercise, 36 (9)*, 1637-1645.

Chapter 3

SOURCES OF VARIABILITY IN DAILY PHYSICAL ACTIVITY FOR MIDDLE SCHOOL STUDENTS

So-Yeun Kim

Abstract

Using generalizability theory, the sources of variability in daily physical activity levels of adolescents and a determination of the minimum number of days required for estimating typical physical activity behaviors were examined. Thirty one (*M* = 12.3 years, *SD* = 0.5 years) middle school students wore two pedometers and two accelerometers during 5 weekdays and 4 weekend days. A two random facet (instrument and number of days) completely crossed design was conducted. Six separate analyses were employed to identify sources of variability for accelerometers and pedometers across weekdays (W), weekends (WK), and weekdays and weekend days combined (WWK). During W and WK, the primary sources of variability were the variance components of the person (49.13% and 32.33%, respectively) and the person by day interaction (48.07% and 66.21%, respectively) for pedometers. Using accelerometers, large amounts of variability were related to the persons (53.90% and 25.92%, respectively), and the person by day interaction component (29.96% and 57.37%, respectively) across W and WK. During WWK, a relatively large percentage of variability was associated with residuals for pedometers (35.23%) and accelerometers (35.59%). The minimum numbers of days required for monitoring using one pedometer were 5 and 9 days during W and WK, respectively. Using one accelerometer, 4 and 14 days of monitoring physical activity were needed during W and WWK, respectively. It was determined that estimation of physical activity levels during WWK using one pedometer and during WK using one accelerometer was problematic within reasonable length of time.

Sources of Variability in Daily Physical Activity for Middle School Students

Considerable effort has been made to determine daily physical activity levels of children in free living settings using different physical activity assessment tools (Duke, Huhman, & Heitzler, 2003; Grunbaum et al., 2002; Janz, Dawson, & Mahoney, 2000; Trost et al., 2002). One of the challenges to capture typical physical activity levels of children is the intra-individual variability (Baranowski & de Moor, 2000; Levin, Jacobs, Ainsworth, Richardson, & Leon, 1999; Rowe, Mahar, Raedeke, & Lore, 2004). Intra-individual variability in the daily physical activity indicates differences in physical activity levels in a person from day to day (Baranowski & de Moor, 2000). Intra-individual variability in the daily physical activity of children has been estimated by identifying the minimum number of days of measurement required to get representative information of typical physical activity in a given population (Rowe et al., 2004; Vincent & Pangrazi, 2002).

The minimum number of days required for monitoring children's typical physical activity has been examined using the intra-class correlation (ICC) with a criterion reliability coefficient of greater than .80 (Baranowski & de Moor, 2000). However, different results have been reported on the optimal number of days required for monitoring typical physical activity levels of children depending on age groups and physical activity instrument used for the studies. When daily physical activity was measured with pedometers, five days for elementary students (Vincent & Pangrazi, 2002) and six days for middle school students (Rowe et al., 2004) were required for monitoring daily physical activity to capture children's typical physical activity levels.

Uniaxial accelerometers have also been used to determine the minimum number of days to monitor moderate and vigorous physical activity (Janz, Witt, & Mahoney, 1995; Trost, Pate, Freedson, Sallis, & Taylor, 2000). Five days of monitoring children (Grade 1-6) and 9 days of monitoring adolescents (Grade 7-12) were suggested (Trost et al., 2000), whereas Janz, Witt, and Mahoney (1995) suggested 6 days of data collection for children between 7 and 15 years old. However, these recommendations assume that intra-individual variability is the only source of variability. In order to gather accurate data of daily physical activity levels for children, both intra-individual variability and inter-instrument reliability should be examined at the same time. Random over-time variability can occur due to both true changes of behaviors within an individual and measurement error (Lakka & Salonen, 1992).

Inter-instrument reliability occurs when physical activity instruments of the same brand provide different data under the same set of circumstances (Bassett & Strath, 2002). Understanding how much variability exists between units and how variable responses are over time is important because it affects the accuracy of data (Welk, Schaben, & Morrow, 2004). There is limited information available on the inter-instrument reliability of daily physical activity in children. Using Yamax Digi-walker pedometers, Barfield, Rowe and Michael (2004) examined inter-instrument reliability measuring the daily physical activity of elementary school students and found excellent reliability coefficient (ICC > .90) between the pedometers.

To date, inter-individual variability and inter-instrument variability have been examined separately. To understand true changes of daily physical activity, other types of error associated with sources of variability should be examined together. Unlike classical test theory, Generalizability theory (G-theory) allows researchers to estimate the magnitude of multiples sources of potential error (called ‘facets’) in one study (Shavelson & Webb, 1991). The estimate of the participants’ variability across time as well as systematic errors in days of measurement and measurement errors due to unsystematic sources can be examined with G-theory (Shavelson & Webb, 1991). Moreover, G-theory can provide optimal measurement protocols that minimize variability for a particular condition (Shavelson, & Webb, 1991; Shavelson, Webb, & Rowley, 1989; Ulrich & Wise, 1984).

G-theory techniques have not been used extensively in the measurement of physical activity studies although researchers in the field of measurement in exercise sciences have recently proposed using G-theory in the measurement of physical activity studies. This fact may be explained by its high demand on time and effort compared to simpler approaches to reliability estimation (Goodwin, 2001).

To my knowledge, there is no research in the measurement of daily physical activity for children that has used G-theory in real life settings. One published study employed G-theory to examine the variability of daily physical activity levels of sedentary adult males (Coleman & Epstein, 1998). Only one facet, ‘days’, was included in this study. In the study, vector magnitude of accelerometer and METs (multiples of resting metabolic rate) calculated with accelerometer data were used. The largest percentage of variation was associated with residual term for vector magnitude (46%), and METs (56%). The day facet was associated with only 1% and 2% of the total variance, respectively. The largest magnitudes of residual terms indicate that other facets, which might be associated with the variability in daily physical activity of sedentary males, were not included in the study. The results indicate that more complicated designs, containing more than one facet, should be used to examine variability of daily physical activity in the participants.

The primary purpose of this study was to examine sources of variability in the measurement of daily physical activity levels of adolescents. The secondary purpose was to investigate minimum number of days required for monitoring daily physical activity levels to determine their typical physical activity of adolescents. The same questions were examined both for pedometers and accelerometers.

Method

Participants

A convenience sample of 31 middle school students (15 boys and 16 girls) in Grade 6 (M = 12.3 years, SD = 0.5 years) was recruited from a middle school in the rural area of Northwestern United States. Informed consent forms were sent to 6th grade students and their parents from the middle school, and students who returned the signed written informed consent forms participated in the study. This study was approved by a University Institution Review Board prior to recruiting participants.

Instruments

Accelerometers

Actiwatches (Mini-Mitter, Bend, OR) were used to measure the daily physical activity levels of the participants. This device has an omnidirectional sensor that is able to monitor acceleration of body movement in all directions, but is most sensitive in the direction parallel to the longest dimension of the case (Puyau, Asolph, Vohra, & Butte, 2002). The speed and motion of body movements can be monitored according to user-defined epochs from 15 seconds to 1 minute (Mini-Mitter, 2000). In this study, 30 seconds intervals were used.

To ensure the consistency among the accelerometers, prior to the data collection, each device was worn at least 10 times; two researchers wore three accelerometers on the waist for 3 hours. Devices, which were produced ±15% of error, were recalibrated by the manufacturer. Then accelerometers were recalibrated. Error was calculated with the following equation;

$$\frac{X_i - [(X_1 + X_2 + X_3 + X_4 + X_5 + X_6)/6]}{X_i} \times 100$$

Note: Where *X* represents the average activity counts of an accelerometer, and *i* represents the sample average.

Pedometers

Omron HJ-112 pedometers (Omron Healthcare, Vernon Hills, 2003) were used in this study. According to the manufacturer (Omron Healthcare, 2003), this pedometer has two unique features. First, the pedometer accumulates steps only after walking at least four seconds. This function can eliminate the chance of counting non step movements as steps. Second, the pedometer is able to count steps correctly when the front of the main unit is placed with an angle of more than 60 degrees from the ground or horizontal to the ground. This feature may have a positive influence on accuracy in individuals with high waist circumference because this pedometer may be less susceptible to errors that occur due to tilt**.** The pedometer can store steps for seven days. It has an internal clock that automatically resets the counts to zero, so users do not have to press the reset button every day.

The "shake-test" developed by Vincent and Sidman (2003) was used to examine the proper calibration of the pedometers. Each pedometer was placed in a shipping box from the manufacturer, and shaken 100 times. To decrease aberrant movements of the pedometers, the box was moved in a vertical direction while the bottom of the box remained in contact with a table surface. These tests were performed twice for each pedometer before and after the study. No pedometers exceeded ± 5% error.

Procedures

All the participants were asked to wear four devices including two pedometers and two accelerometers for 11 days including 7 weekdays and 4 weekend days between April and June. The first 2 weekdays were for familiarizing participants with wearing the devices and for decreasing reactivity that the measurement process might influence the participants' physical behaviors. Randomly selected two pedometers and two accelerometers were given to each participant, and the devices were replaced every other day unless students missed school. Also, participants were asked to keep a daily physical activity log, so that the researcher would have detailed information about the self-reported types of physical activity. Physical activity logs were collected daily. Weekend logs were colleted on Mondays.

Participants were instructed to wear two pedometers attached to a Velcro belt on their waists, in line with the middle of their right thigh, from the time that they got up in the morning until bedtime except while swimming and showering since pedometers were not waterproof. Wearing pedometers with a Velcro belt helped participants to wear them in the correct place. Participants were asked to take off the Velcro belt, not the pedometers when they had to take off their pedometers, so they would wear them consistently in the correct place. In order to prevent inflating counts, all participants were instructed not to tamper with the instruments. Moreover, pedometers were sealed to prevent accidentally resetting the instruments to zero.

To remind participants to wear the pedometers every day, the researcher gave two reminder posters to each participant. Participants were asked to attach the posters wherever they could be easily seen. Also, a reminder phone call was made by the researcher every morning to 21 participants, or their parents who wished to receive phone calls. All participants were also asked to wear two accelerometers on their wrist of non- dominant hand for 24 hours. To ensure wearing them all day, two accelerometers attached by a non-removable plastic wrist band attached to each participant. Moreover, participants were instructed not to cut the band unless they felt discomfort.

Data Analysis

G-theory was used in this study. Two types of studies can be conducted using G-theory, including generalizability study (G-study) and decision study (D-study) (Morrow, 1989). A G-study is designed to provide estimates of the variance components associated with each facet and their interactions while a D-study is designed to make substantive decisions about a measurement protocol using the results of the G-study (Morrow, 1989; Shavelson, Webb, & Rowley, 1989).

<u>G-study</u>

A two-facet fully crossed design was employed to answer the first research questions about the sources of variability of daily physical activity. Both instruments and days were considered random facets. A facet is defined as sources of error (Shavelson, Webb, & Rowley, 1989). Fully crossed designs are defined as those in which all of the facets are treated as random, and all participants are crossed with all other facets in the model (Morrow, 1989). If facets are considered exchangeable with others facets, the facets are treated as random (Shavelson & Webb, 1991).

Pedometers and accelerometers cannot be exchanged because they measure different aspects of physical activity behaviors. Thus, the data were analyzed separately. Weekdays and weekends may not be interchangeable because physical activity patterns during weekdays and weekend days may be different. However, limited information is available on differences in physical activity patterns of children on weekdays and weekends. The data were analyzed in three different ways; weekdays, weekends, and combined weekdays and weekend days.

Six two-way ANOVAs were used to examine the sources of variability in daily physical activity levels. Average steps per day and the mean of total activity scores per day were used for pedometers and accelerometers, respectively. Seven variance components were estimated using the VARCOMP procedure from SAS. The seven sources of variability include variance associated with persons (σ^2_p), days (σ^2_d), and instruments (σ^2_i), three two-way interaction including persons by days (σ^2_{pd}), persons by instruments (σ^2_{pi}), and instruments by day (σ^2_{id}), and the residual term (three-way interaction plus error) ($\sigma^2_{pid,e}$). Negative variance components were set to zero in subsequent calculations as suggested by (Morrow, 1989). The percentage of variance associated with each source of variability was calculated by dividing each variance estimate by the total variance.

Generalizability coefficients, G and phi coefficients were also calculated from the estimated variance components. G coefficients are calculated based on the relative decision while phi coefficients are calculated based on absolute decisions (Shavelson, Webb, & Rowley, 1989). Relative decisions refer to decisions that are based on the relative standing of individuals whereas absolute decisions are based on the absolute level of performance. In relative decisions, variance components of person by instrument, person by day, and residual term contribute to error while all variance components except person are considered as error in absolute decisions (Shavelson, Webb, & Rowley, 1989). The following formula was used (Shavelson & Webb 1991);

$$G = \sigma_p^2 / (\sigma_p^2 + \sigma_{pd}^2 / n_d + \sigma_{pi}^2 / n_i + \sigma_{pid,e}^2 / n_i \; n_d)$$

$$phi = \sigma_p^2 / (\sigma_p^2 + \sigma_i^2 / n_i + \sigma_d^2 / n_d + \sigma_{pd}^2 / n_d + \sigma_{pi}^2 / n_i + \sigma_{id}^2 / n_i \; n_d + \sigma_{pid,e}^2 / n_i \; n_d)$$

D-study

The second research questions, the optimal number of days required for monitoring daily physical activity levels was examined using a D-study. By increasing or decreasing the number of instruments and days, the minimum number of facet levels required to establish the desired generalizability was assessed.

Results

Table 3.1 presents descriptive statistics for each device across weekdays, weekends, and weekdays and weekend days. Only participants with complete data of pedometers and accelerometers across weekdays, weekends, and weekdays and weekend combined were used. The average steps per weekday, weekend day, and weekday and weekend days were 9432, 6344, and 7933 steps, respectively. The average activity counts per weekday, weekend day, and weekday and weekend days measured with accelerometers were 541438, 473199, and 510396 activity counts, respectively.

Table 3.1. Descriptive statistics for pedometers and accelerometers

	Pedometer			Accelerometers		
	N	Mean	SD	N	Mean	SD
Weekdays	31	9,432	3,036	30	541,438	125,406
Weekends	30	6,344	2,414	29	473,199	138,633
Weekdays/Weekends	30	7,933	2,851	29	510,396	135,493

Accelerometers

G-study

Table 3.2 presents the estimated variance components and relative magnitude for accelerometers during weekdays, weekend days, and weekdays and weekends combined. During weekdays, the primary sources of variability were the person (53.90%) and the person by day interaction (29.96%). Relatively large variance components for residual term (12.76%) were estimated. The magnitude for other sources of variability was negligible. The estimated generalizablity coefficients for the average of two accelerometer data across five weekdays was high, *phi* = .87, *G* = .88.

During weekend days, the largest percentage of variability was associated with the person by day interaction component (57.37%). Relatively large variance was due to the persons (25.92%), and relatively low variances were related to days (5.27%), the person by instrument (5.25%), and the residual term (6.17%). Variability due to the instruments facet, and the instruments by day interaction components were very low. The estimated generalizability coefficients for the average of two accelerometer data across four weekend days was moderate, *phi* = .58, *G* = .59.

When data from weekdays and weekends were combined, relatively large sources of variability were associated with the persons (38.61%) and residual term (35.59%). The person by day interactions was associated with 12.75% of the total variance while the variance associated with the instruments by days was 7.31%. Negligible variances were related to instruments (2.02%) and day facets (0%), and the person by instruments interaction component (3.71%). The estimated generalizability coefficients for the average of two accelerometer data across nine days including weekdays and weekends was high, *phi* = .85, *G* = .88.

Table 3.2. Variance Component Estimates and Relative Magnitudes for Accelerometers

Variation	Estimated Variance Components	Relative Magnitude[1]
Weekdays		
Persons (p)	8695602568	53.90%
Instruments (i)	132096845	0.82%
Days (d)	301549300	1.87%
p × i	107116689	0.66%
p × d	4833301422	29.96%
i × d	3609564.4	0.02%
Residual (p × i × d, *e)*	2058351233	12.76%
Weekends		
Persons (p)	5110187427	25.92%
Instruments (i)	-47278646[2]	0.00%
Days (d)	1038263248	5.27%
p × i	1035843894	5.25%
p × d	11311800000	57.37%
i × d	4290582.6	0.02%
Residual (p × i × d, *e)*	1216539213	6.17%
Weekdays/Weekends		
Persons (p)	6790281754	38.61%
Instruments (i)	355602309	2.02%
Days (d)	-426408991[2]	0.00%
p × i	652811813	3.71%
p × d	2242646572	12.75%
i × d	1285199988	7.31%
Residual (p × i × d, *e)*	6258443819	35.59%

Note. [1] Relative magnitude was calculated using estimated variance divided by the total variance. [2] Negative variance components were set to zero in subsequent calculations as suggested by Morrow (1989).

D-study

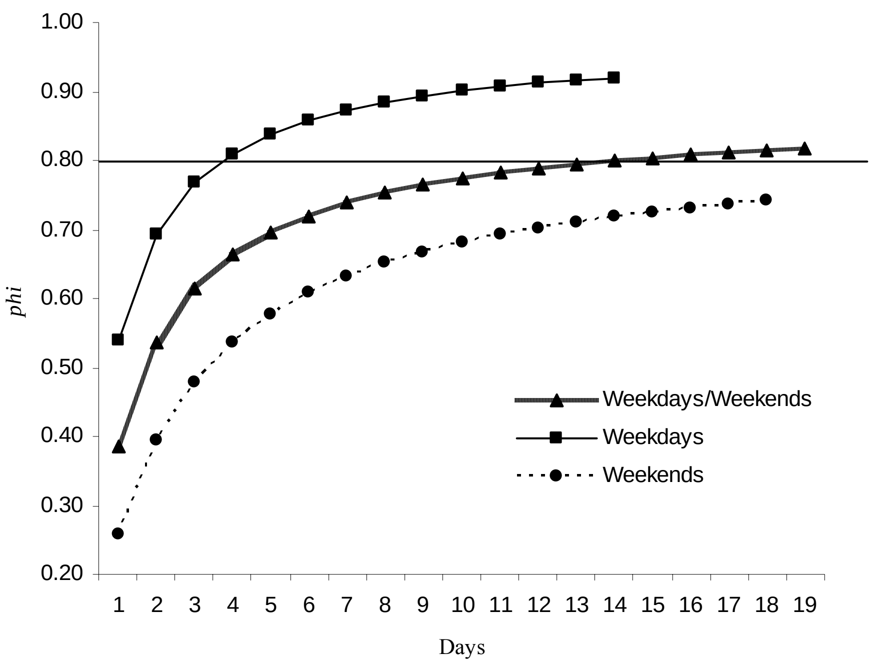
The optimal number of days required for monitoring daily physical activity was determined with the D-study (Table 3.3). In practical settings, it is unlikely that professionals will use more than two accelerometers to measure daily physical activity of children. Thus, generalizability coefficients were estimated when one or two accelerometers were used.

To achieve sufficient reliability (*phi* = .80, *G* = .80) in the measurement of daily physical activity during weekdays, measurements need to be obtained on at least 4 days using an accelerometers, or 3 days using two accelerometers. To determine daily physical activity levels of children during weekend days, much longer periods of measurement was estimated. At least 19 weekend days using two accelerometers, or 50 weekend days using one accelerometer were estimated to capture typical weekend physical activity levels of the participants. When weekdays and weekend days were combined, measurements should be obtained from 14 days with one accelerometer or five days with two accelerometers. Estimated phi coefficients across days using an accelerometer during weekdays, and weekdays and weekends combined are presented in Figure 3.1 shows estimated phi coefficients as the number of measurement days were increased using one accelerometer.

Table 3.3. Estimated number of monitoring days using Accelerometers

Number of Days	Number of Instruments	*phi*	*G*
Weekdays			
3	2	.80	.81
4	1	.81	.83
5	1	.84	.85
Weekends			
8	2	.58	.72
19	2	.81	.82
50	1	.80	.80
Weekdays and Weekends			
14	1	.80	.84
5	2	.80	.83
7	1	.74	.84

Figure 3.1. Estimated phi coefficients across days using one accelerometer



Pedometers

G-study

Estimated variance components and relative magnitude for accelerometers during weekdays, weekends, and weekdays and weekend days combined are presented in Table 3.4. During weekdays, the primary sources of variability were the person (49.13%) and the person by day interaction (48.07%). The magnitude for other estimated variance components were negligible. The estimated generalizability coefficient for the average of two pedometer data across five weekdays was high, *G* = .83, *phi* = .83.

During weekends, the largest percentage of variability was associated with the person by day interaction component (66.21%) while the variance associated with the persons was 32.33%. Again, the magnitude for other estimated variance components were negligible. The estimated generalizability coefficient for the average of two pedometer data across four weekend days was moderate, *G* = .66, *phi* = .65.

When physical activity data obtained from weekdays and weekends were combined, the largest percentage of variability was associated with the residual term (35.23%). Relatively large variance was due to the persons (25.13%) and the persons by days (16.93%) components. Also, the variance associated with the instruments facet, the instruments by day interaction, the persons by instruments interaction were 7.18%, 9.80%, and 4.98%, respectively. The generalizability coefficients, *G* = .80, *phi* =.70, were estimated when 9 days of daily physical activity were measured with 2 pedometers.

Table 3.4. Variance Component Estimates and Relative Magnitudes for Pedometers

Variation	Estimated Variance Components	Relative Magnitude[1]
Weekdays		
Persons (p)	4608384	49.13%
Instruments (i)	-391.3[2]	0.00%
Days (d)	55071.4	0.59%
p × i	39502.7	0.42%
p × d	4508880	48.07%
i × d	1566.8	0.02%
Residual (p × i × d, *e)*	166355.3	1.77%
Weekends		
Persons (p)	3820066.6	32.33%
Instruments (i)	6576.7	0.06%
Days (d)	80477.8	0.68%
p × i	73066.8	0.62%
p × d	7823793.2	66.21%
i × d	1343.2	0.01%
Residual (p × i × d, *e)*	11043.3	0.09%
Weekdays/Weekends		
Persons (p)	3060364.4	25.13%
Instruments (i)	874260.8	7.18%
Days (d)	91772.7	0.75%
p × i	606530.5	4.98%
p × d	2061636.4	16.93%
i × d	1193603.0	9.80%
Residual (p × i × d, *e)*	4290172.8	35.23%

Note. [1] Relative magnitude was calculated using estimated variance divided by the total variance. [2] Negative variance components were set to zero in subsequent calculations as suggested by Morrow (1989).

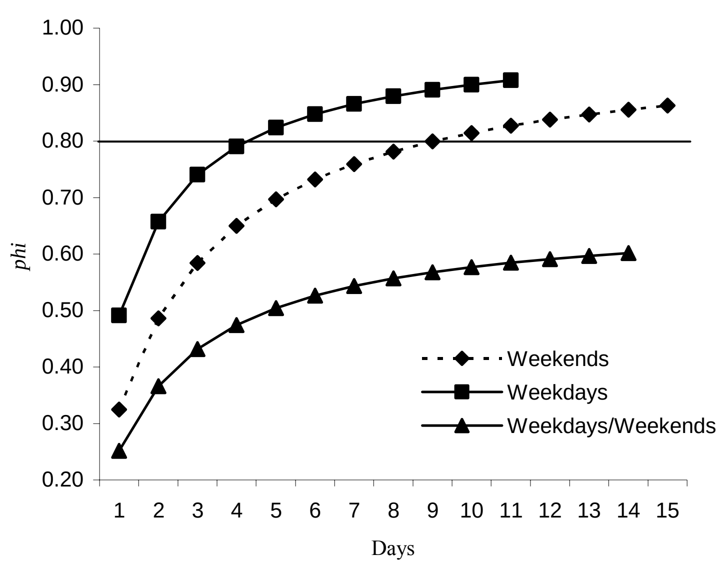
D-study

Table 3.5 presents the results of D-study. To achieve sufficient reliability in the measurement of daily physical activity during weekdays, measurements need to be obtained on at least 5 days using a pedometer, or 4 days using two pedometers. To determine typical physical activity levels of children during weekend days, at least 9 weekend days of measurement using a pedometer was estimated to be required. When data from weekdays and weekends were combined, at least 14 days of measurements using 3 pedometers were needed. Estimated phi coefficients as the number of measurement days were increased using one pedometer are presented in Figure 3.2.

Table 3.5. Estimated days of observation using pedometers

Number of Days	Number of Instruments	*phi*	*G*
Weekdays			
3	1	.74	.74
4	2	.80	.80
5	1	.82	.83
Weekends			
10	1	.81	.82
9	2	.81	.81
9	1	.80	.80
Weekdays and Weekends			
60	1	.66	.81
30	2	.77	.87
14	3	.80	.87

Figure 3.2. Estimated phi coefficients across days using one pedometer



Discussion

Accelerometers

The results of G-study for accelerometers determined that the primary sources of variability in daily physical activity levels of children during weekdays were the variance component of persons (53.90%) and persons by day interaction (29.96%). The large variance component of persons indicates that the variability was related to true differences of physical activity levels among the participants. The large person by day interaction component suggests that some participants were very active on particular days, but not other days. Participating in sports on certain days during the week could be reasons for the large person by day interaction. According to physical activity logs, some participants regularly participated in a certain physical activity, two or three times a week. Also, during physical education classes, participants often had to decide on physical activity from several activity choices including archery, track and field, basketball, baseball, rock climbing, and weight training. Thus, depending on what kind of physical activity they chose during the physical education, some participants might have higher activity counts on some days.

Amounts of variability were related to the instruments were negligible. Variability associated with instruments was 0.82%, indicating very small amount of systematic error was associated with instrument. Also, only a very small amount of variability was associated with the persons by instrument interaction component (0.66%). This indicates that the instruments produced consistent activity counts across the participants.

During weekends, variability associated with differences among participants was 25.92% while the person by day interaction component was associated with 57.37% of the total variability. The large variance of the person by day interaction suggested that some students were active on a particular day, but not another. One of the reasons for this may be that there was not much routine activity during weekends. According to the physical activity logs, seven students participated in team sports every Saturday. When participants visited parks or beach with their parents, or engaged in physical activity in the outside they did get many counts. Also, several students mentioned that they stayed at home and played computer games during weekends.

When data from weekdays and weekends were combined, different results were found. The largest percentage of variability came from the residual components (35.59%). This indicates that a substantial amount of the variance was related to the three-way interaction (p × i × d) and to unexplained sources of error not measured in the study. This variance could be due to different physical activity patterns between weekdays and weekends, which indicate physical activity levels during weekdays and weekends should be treated differently. Therefore, physical activity during weekdays should not be interchangeable with physical activity during weekends. Another reason of relatively large amount of variance in residual component could be weather. Several students mentioned that they did not engage in physical activity due to hot weather or rain.

Relatively large percentages of variation were associated with the person (38.61%) and the person by day interaction (12.75%) when weekdays and weekends were combined. Also, small amounts of variability were related to the instrument facet (2.02%), and the instrument by day (7.31%), and the person by instrument (3.71%) components. The variance component of instrument by day interaction indicates an instrument inconsistently measures physical activity from one day to another day. The person by instrument interaction indicates inconsistencies in activity counts between instruments for particular persons.

The minimum number of days required for monitoring daily physical activity levels was determined with the D-study based on the result of the G-study. Estimated phi coefficients were only discussed because relative standing of participants in this study had no bearing on the results; the physical activity of each person did not depend on the physical activity of his or her peers. Results of D-study indicate that to achieve sufficient reliability (*phi* $\geq$.80) for determining the typical physical activity levels of children on weekdays, measurements need to be obtained on at least 4 days using one accelerometer, or 3 days using two accelerometers. To capture typical physical activity levels of adolescents during weekends, at least 19 weekend days must be monitored using two accelerometers, or 50 weekend days must be monitored if only one accelerometer were used. During weekend days, there was much greater variability related to the person by day interaction (57.37%) in daily physical activity than the persons (25.92%), resulting the estimation of many more number of days required for monitoring physical activity to achieve reliability of .80. Using multiple measures of daily physical activity to determine typical physical activity levels of adolescents during weekend days may reduce intra-individual variability in the assessment of physical activity (Baranowski & de Moor, 2000).

When weekdays and weekend days were combined, the minimum number of days was estimated at 14 days of monitoring with one accelerometer or 5 days with two accelerometers. No published study has examined the minimum number of days required to monitor typical physical activity with adolescents using the G-theory. However, different recommendations have been proposed for determining moderate and vigorous physical activity levels of adolescents without disabilities using the intra-class correlation. Using uni-axial accelerometers, Trost et al. (2000) recommended 5 days of monitoring children (Grade 1-6) and 9 days of monitoring adolescents (Grade 7-12), whereas Janz, Witt, and Mahoney (1995) suggested 6 days of data collection for children and adolescents between 7 and 15 years old. There were two issues in the previous studies. First, those studies examined only one source of variability (day), and did not count variability related to instruments. Second, data from physical activity during weekdays and weekend days were not analyzed separately. According to the results of this study, physical activity patterns of middle school students might be different during weekdays and weekend days. Also, Durant et al. (1993) examined compound symmetry of physical activity data from children for three consecutive weeks using 7-day physical activity records. Compound symmetry means that variance and covariance are equal across units (Baranowski & de Moor, 2000). Therefore, each day can be replaced by a different day. Large estimated variance components in three way interaction and random components after week and weekday combined may suggest that the week and weekends may not be mixed while capturing middle school students' physical activity levels.

Pedometers

The results of G-study for pedometers were similar to the results of accelerometers. During weekdays, the primary sources of variability were the variance component of persons (49.13%) and persons by day interaction (48.07%), indicating large differences in the daily physical activity levels among the participants, and some participants were physical active on particular days, but not the other days. Negligible amounts of variance were associated with other variance components. These results indicate that variability in daily physical activity of adolescents during weekends was associated with true physical activity changes rather than measurement errors.

During weekends, the primary sources of variability were also related to the persons (32.33%) and the person by day interactions (66.21%). Again, relatively large proportion of variance of daily physical activity levels were associated with different physical activity levels among the participants. Also, some participants were physically active during some weekend days, but not others. Negligible amount of variance was related to other variance components (less than 1%).

When data from weekdays and weekend days were combined, large residual components were associated with 35.23% of the total variance. This result can be explained by different physical activity patterns during weekdays and weekends. Data collected during weekdays and weekends might not be exchangeable. Therefore, they should be analyzed separately. A relatively large percentage of the variation was associated with the person (25.13%) and the person by day interaction (16.93%). In addition, small amounts of variability were related the instrument facet (7.18%), and the instrument by day (9.80%), and the person by instrument (4.98%) components. The results indicate that low inter-instrument reliability, and consistent measurement across days and persons.

The results of D-study indicate that at least 5 days of monitoring daily physical activity levels (*phi* = .82) using one pedometer, or 4 days using two pedometers (*phi* = .80) was required to determine the typical physical activity levels of children during weekdays. To determine daily physical activity levels of children during weekends, at least 9 weekend days of monitoring physical activity using one pedometer (*phi* = .80) were estimated to be needed. The same results were found when two pedometers were used to measure daily physical activity for 9 weekend days (*phi* = .80). When weekdays and weekend days were combined, the minimum number of days required for monitoring daily physical activity was at 14 days with 3 pedometers. Because the variability related to instrument was 7.18%, using more pedometers could reduce the variability sensationally. Therefore, this approach may be preferable to increasing the number of days of monitoring.

The minimum number of days required to monitor typical physical activity of children using the G-theory has not been examined empirically. However, recommendations on the minimum number of days required for monitoring have been provided using ICC based on the assumptions that physical activity patterns during weekdays and weekends are same (Rowe et al., 2004; Vincent & Pangrazi, 2002). Using intraclass correlation coefficient (*R*), some recommendations have been made: Six days (R = .83) of monitoring for middle school students (Rowe et al., 2004) and 5 days (R = .82) of monitoring for elementary students (Vincent & Pangrazi, 2002). One of the problems in the previous studies was that the previous studies did not assess measurement errors related to the instruments. Also, data from physical activity during weekdays and weekend days were not analyzed separately. The results of this study, the relatively large residual component (35.59%) indicate that physical activity patterns during weekdays and weekend days may be different. Durant et al. (1993) examined compound symmetry of physical activity data of children during weekdays and weekends using 7-day physical activity, and found significant differences between weekday and weekend physical activity patterns. Therefore, physical activity during weekdays and weekends should be analyzed separately.

Summary

Variability in daily physical activity levels of middle school students was associated with behavioral variability rather than analytic variability. These results indicate that profiles of different physical behavior patterns for middle school students should be identified, especially during weekends. Why middle school students' physical activity levels vary across days, and why some students are more activity some days, but not on others should be examined. The information may help not only to decide units of measurement in physical activity but also to develop effective intervention strategies to promote daily physical activity levels for middle school students.

The relatively large residual components for both pedometers and accelerometers during weekdays and weekends combined indicate that physical activity patterns of middle school students during weekdays and weekends might be different. Further research should examining different daily physical activity patterns during weekdays and weekends.

These data indicate that it is difficult to estimate typical physical activity behaviors of middle school students when weekend days are included. It is also unrealistic to use multiple instruments over multiple days to get a sufficient level of reliability. Further investigation is needed to analyze the contextual framework of students who have more reliable estimates over weekends in order to develop programs to increase physical activity behaviors.

References

Baranowski, T., & de Moor, C. (2000). How many days was that? Intra-individual variability and physical activity assessment. *Research Quarterly for Exercise and Sport, 71*(2), 74-78.

Barfield, J. P., Rowe, D. A., & Michael T. J. (2004). Interinstrument consistency of the Yamax Digi-Walker pedometer in elementary school-aged children. *Measurement in Physical Education and Exercise Science, 8*(2), 109-116.

Bassett, D. R., & Strath, S. J. (2002). Use of pedometers to assess physical activity. In G.J. Welk (Ed.), *Physical activity assessments for health-related research (pp.163-177)*. Champaign, IL.

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

Duke, J., Huhman, M., & Heitzler, C. (2003). Physical activity levels among children aged 9 to 13 years: United States, 2002. *Morbidity and Mortality Weekly Reports, 52*(33), 785-788.

Durant, R. H., Baranowski, T., Puhl, J., Rhodes, T., Davis, H., Greaves, K. A., & Thompson, W. O. (1993). Evaluation of the Children's Activity Rating Scale (CARS) in young children. *Medicine and Science in Sports and Exercise*, *25,* 1415-1421.

Goodwin, L.D. (2001). Interrater agreement and reliability. *Measurement in Physical Education and Exercise Science, 5(1),* 13-34.

Grunbaum, J. A., Kann, L., Kinchen, S., Williams, B., Ross, J. G., Lowry, R., et al.(2002). Youth risk behavior surveillance-United States, 2001. *Morbidity and Mortality Weekly Report, 51*(SS-4), 1-64.

Janz, K. F., Dawson, J. D., & Mahoney, L. T. (2000). Tracking physical fitness and physical activity from childhood to adolescence: The Muscatine study. *Medicine and Science in Sports and Exercise*, *32*(7), 1250-1257.

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

Lakka, T. A., & Salonen, J. T. (1992). Intra-person variability of various physical activity assessments in the Kuopio ischemic heart disease risk factor study.

International Journal of Epidemiology, *21*(3), 467-472.

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

Minimitter Company Inc. (2000). *Actiwatch activity monitors: Instruction manual*. Sunriver, OR: Author.

Morrow, J. R. (1989). Generalizability theory. In M.J. Safrit & T.M. Wood (Eds.), *Measurement concepts in physical education and exercise science (pp.73-96).* Champaign, IL: Human Kinetics.

Omron HealthCare Inc. (2003). Omron instruction manual: Pedometer Model HJ-112.Retrieved April 14, 2006, from http://www.omronhealthcare.com/enTouchCMS/app/viewDocument?docID=1531&parntCatgId=34

Puyau, M.R., Adolph, A.L., Vohra, F.A., & Butte, N.F. (2002). Validation and calibration of physical activity monitors in children. *Obesity Research, 10*(3), 150-157.

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

Shavelson, R. J., & Webb, N. M. (1991). *Generalisability Theory: A primer*. Newbury Park, CA: Sage.

Shavelson, R. J., Webb, N. M., & Rowley, G. L. (1989). Generalizability theory. *American Psychologist, 44*(6), 922-932.

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

Trost, S. G., Pate, P. R., Sallis, J. F., Freedson, P. S., Taylor, W. C., Dowda, M., et al. (2002). Age and gender differences in objectively measured physical activity in youth. *Medicine and Science in Sports and Exercise*, *34*(2), 350-355.

Ulrich, D. A., & Wise, S. L. (1984). Reliability of scores obtained with the objectives based motor skill assessment instrument. *Adapted Physical Activity Quarterly, 1*, 230-239.

Vincent, S. D., & Sidman, C. L. (2003). Determining measurement error in digital pedometers. *Measurement in Physical Education and Exercise Science, 7*, 19-24.

Vincent, S. D., & Pangrazi, R. P. (2002). Does reactivity exit in children when measuring activity levels with pedometers? *Pediatric Exercise Science, 14,* 56-63.

Welk, G. J., Schaben, J. A., & Morrow, J. R. (2004). Reliability of accelerometry-based activity monitors: A generalizabilty study. *Medicine and Science in Sports and Exercise, 36*(9), 1637-1645.

Chapter 4: General Conclusions

The following sections are included in 1) overall conclusions for the two studies including further research directions and 2) conclusion of each research question presented in the general introduction.

Overall Conclusions

The main sources of variability in daily physical activity levels of secondary students with and without DD were found to be associated with behavioral variability rather than analytic variability. Possible questions to be explored in further studies are 1) why some children's physical activity levels are high on some days, but not on others, 2) why some children have consistent physically activity across days, but not others, 3) who are more consistently physically active during weekends, and 4) what are the contextual factors of those who are highly active and those who are not.

The relatively low variance related to instruments suggests high reliability evidence of Omron HJ-112 pedometers and Actiwatches for the measurement of physical activity in secondary students with and without DD during free living settings. Further research should examine validity evidence of both devices in daily physical activity measurement during free living settings. Especially, accuracy of Omron HJ-112 pedometers on individuals with obesity should be examined further. Inaccuracy of pedometers for individuals with obesity has been reported (Shepherd, Toloza, McClung, & Schmalzried, 1999). However, relatively low variance of the person by instrument component for the pedometers might indicate that the pedometer accurately measure steps of individual across different body shapes. This could be due to the unique function of the pedometer, which is the ability to count steps correctly when the front of the main unit is placed at an angle of more than 60 degrees from the ground or horizontal to the ground. Because of this function, Omron HJ-112 pedometers might be less susceptible to errors that occur due to tilt. Future research should investigate further the accuracy of this pedometer for individuals with obesity.

Recommendations on the minimum number of days required for monitoring have been provided based on the assumption that physical activity patterns during weekdays and weekends are same (Janz, Witt, & Mahoney, 1995; Rowe et al., 2004; Trost et al., 2000; Vincent & Pangrazi, 2002). However, the relatively large residual component for pedometers (35.23%) and accelerometers (35.59%) for students without disabilities indicate that physical activity patterns during weekdays and weekend days may be different. Compound symmetry of physical activity during weekdays and weekend days should be tested using daily physical activity data with pedometers and/or accelerometers.

To decrease the minimum number of days monitoring necessary, especially during weekends, different units of measurement of daily physical activity should be considered. Monitoring only during the day from in the morning to afternoon, when children are expected to be physically active, might produce less variability than measurement throughout the day.

Research Questions and Conclusions

1. How much of the variance in daily physical activity levels in secondary students with DD and without disabilities was related to the days?

<u>Secondary Students with DD</u>

When physical activity was assessed with accelerometers, relatively large variability was associated with the person by day interaction (25.76%, 38.04%, and 37.15%, respectively) while very little variability was related the day facet (0%, 0%, and 4.11%, respectively) across weekdays, weekends, and weekdays and weekend days combined. When physical activity was measured with pedometers, relatively large variability was associated with the person by day interaction (44.04% and 57.95%, respectively) while very little variability was associated with the day facet (0% and 0%, respectively) during weekdays and weekends. However, during weekdays and weekend days combined, the day facet was related to 13.71% of total variance while the person by day interaction was associated with 52.61% of the total variance.

The results indicate that some people accumulate variable amounts of activity counts on particular days. Also, there were a few days that all the participants had increased or decreased steps or activity counts. These results indicate systematic interaction between person and day although participants had similar routine environment.

<u>Secondary Students without Disabilities</u>

During weekdays and weekend days, relatively large variance was due to the person by day interaction for accelerometers (29.96% and 57.37%, respectively) and for pedometers (48.07% and 66.21%, respectively) while relatively small variance was associated with the day facet for accelerometers (1.87% and 5.27%, respectively) and for pedometers (0.59% and 0.68%, respectively). However, when data from weekdays and weekend days were combined, the small amount of sources of variability was due to the person by day interaction for accelerometers (12.75%) and for pedometers (16.93%). Moreover, negligible variances were related to the day facet for accelerometers (0%) and pedometers (0.75%). The results indicate that there was no single day, which particularly required more or less activity counts or steps from the participants.

2. How much of the variance in daily physical activity levels in secondary students with DD and without disabilities was related to the instruments?

<u>Secondary Students with DD</u>

Negligible variances of the instrument facet, the person by instrument interaction, and instrument by day interaction were estimated when data were analyzed in three different way including weekdays, weekends, and weekdays and weekend days combined. All of the variance components were less than 0.2% for accelerometers and pedometers. The results indicate consistent activity counts were produced between accelerometers, no inconsistent activity counts between accelerometers for particular persons, and consistent activity counts between accelerometers from one day to another day. These results provide high reliability evidence of Omron HJ pedometers and Actiwatches for the measurement of physical activity in secondary students with DD during free living settings.

Secondary Students without Disabilities

During weekdays, negligible variances of the instrument facet, the person by instrument interaction, and instrument by day interaction were estimated for accelerometers (0.82%, 0.66%, and 0.02%, respectively). During weekend days, the person by instrument interaction was associated with 5.25% of the total variance while the instrument facet (0%) and the instrument by day interaction (0.02%) were negligible.

When physical activity was measured with pedometers, variance components of the instrument facet, the person by instrument interaction, and instrument by day interaction were less than 0.7% during weekdays and weekend days. When data of weekdays and weekend days were combined, more variability, but still relatively low, were due to the instrument facet, the person by instrument component, and the interment by day component for accelerometers (2.02%, 3.71%, and 7.31%, respectively) and for pedometers (7.18%, 4.98%, and 9.80%, respectively).

Relatively low variance related to inconsistent data between devices, and inconsistent data between device for particular persons and from one day to another day indicate that Omron HJ pedometers and Actiwatches have high reliability evidence to measure daily physical activity levels of middle school students.

3. How many days of monitoring physical activity were required to determine habitual or typical physical activity levels of secondary students with DD and

without disabilities using pedometers and accelerometers?

Secondary Students with DD

To capture typical physical activity levels of secondary students with DD using one accelerometer during weekday, weekends, and weekdays and weekend days combined, four days of measurements were required. When physical activity was measured with one pedometer, measurement need to be obtained on at least 4 days for weekday, 7 days for weekend days, and 8 days for weekdays and weekend combined.

Secondary Students without Disabilities

To achieve sufficient reliability in the measurement of daily physical activity using one accelerometer during weekdays, measurements need to be obtained on at least 4 days. However, to determine typical physical activity levels of children during weekends, much longer time was required. At least 50 weekend days of measurement using one accelerometer, or 19 weekend days of measurement using two accelerometers were needed.

Using one pedometer, at least 5 days of measurement were required to determine typical physical activity levels of children during weekdays. During weekends, at least 9 days of measurement should be obtained. When data from weekdays and weekends were combined, much longer time of measurements was required. At least 100 days of measurements using two pedometers, or 14 days of measurements using three pedometers were needed. The results indicate that physical activity patterns during weekdays and weekends may not be same, so data from weekdays and weekend days should not be combined.

Bibliography

American Heritage (2002). The American Heritage Stedman's medical dictionary (2nd Ed.). Retrieved from http://dictionary.reference.com/browse/biological

Baranowski, T., & de Moor, C. (2000). How many days was that? Intra-individual variability and physical activity assessment. *Research Quarterly for Exercise and Sport, 71*(2), 74-78.

Barfield, J. P., Rowe, D. A., & Michael T. J. (2004). Interinstrument consistency of the Yamax igi-Walker pedometer in elementary school-aged children. *Measurement in Physical Education and Exercise Science, 8*(2), 109-116.

Bassett, D. R., Ainsworth, B. E., Leggett, S. R., Mathien, C. A., Main, J. A., Hunter, D. C., & Duncan, G. E. (1996). Accuracy of five electronic pedometers for measuring distance walked. *Medicine and Science in Sports and Exercise, 28,* 1071-1077.

Bassett, D. R., & Strath, S .J. (2002). Use of pedometers to assess physical activity. In G. J. Welk (Ed.), *Physical activity assessments for health-related research (pp.163-177).* Champaign, IL: Human Kinetics.

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.

Blair, S. N. (1993). 1993 C. H. McCloy research lecture: Physical activity, physical activity, and health. *Research Quarterly for Exercise and Sport, 64,* 365-376.

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

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

Cavill, N., Biddle, S., & Sallis, J. F. (2001). Health enhancing physical activity for young people: Statement of the United Kingdom expert consensus conference. *Pediatric Exercise Science, 13*, 12-25.

Chen, K. Y., Cara, S. A., Majchrzak, K., Donahue, C. L., Baker, L., Clemens, L., Sun, M., & Buchowski, M. S. (2003). Predicting energy expenditure of physical activity using hip- and wrist-worn accelerometers. *Diabetes Technology and Therapeutics, 5*(6), 1023-1033.

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

Council for Physical Education for Children (1998). Physical Activity Guidelines for Children: A statement of guidelines. Reston, VA, 1-21.

Corbin, C. B., Pangrazi, R. P., & Franks, B. D. (2000). Definitions: Health, fitness, and physical activity. *The President's Council on Physical Fitness and Sports Research Digest*, Retrieved Augst 26, 2005, from http://www.fitness.gov/digest_mar2000.htm

Dale, D. D, Welk, G. J., & Matthews, C. E. (2002). Methods for assessing physical activity and challenges for research. In G. J. Welk (Ed.), *Physical activity assessments for health-related research (pp.19-34)*. Champaign, IL: Human Kinetics

Department of Health (1995). The Health Survey for England 1993. London: HMSO. In Robertson, A., et al., (2000). Lifestyle related risk factors for poor health in residential settings for people with intellectual disabilities. *Research in Developmental Disabilities, 21*, 469-486.

Dishman, R. K., Washburn, R. A., & Scholeller, D. A. (2001). Measurement of physical activity. *Quest, 53*, 295-309.

Draheim, C. C, Williamsm, D. P., & McCubbin J. A. (2002). Prevalence of physical inactivity and recommended physical activity in community-based adults with mental retardation. *Mental Retardation, 40*(6), 436-444.

Duke, J., Huhman, M., & Heitzler, C. (2003). Physical activity levels among children aged 9 to 13 years: United States, 2002. *Morbidity and Mortality Weekly Reports, 52(33)*, 785-788.

Durant, R. H., Baranowski, T., Puhl, J., Rhodes, T., Davis, H., Greaves, K. A., & Thompson, W. O. (1993). Evaluation of the Children's Activity Rating Scale (CARS) in young children. *Medicine and Science in Sports and Exercise*, *25*, 1415-1421.

Eaton, W., Enns, L., & Presse, M (1987). Scheme for observing activity level: reliability and convergent validity. *Journal of Psychoeducational Assessment, 3*, 273-280.

Eiholzer, U., Nordmann, Y., L'Allemand D., Schlumpf, M., Schmid, S., & Kromeyer-Hauschild, K. (2003). Improving body composition and physical activity in Prader-Willi syndrome. *The Journal of Pediatrics, 142*, 73-78.

Eston, R. G., Rowlands, A. V., & Ingledew, D. K. (1998). Validity of heart rate, pedometry, and accelerometry for predicting the energy cost of children's activities. *Journal of Applied Physiology, 84*, 362-371.

Faison-Hodge, J., & Porretta, D. L. (2004). Physical activity levels of students with mental retardation and students without disabilities. *Adapted Physical Activity Quarterly, 21*, 139-152.

Fernhall, B. (2002). Physical activity, metabolic issues, and assessment. *Physical Medicine and Rehabilitation Clinics of North America, 13,* 925-947.

Goodwin, L. D. (2001). Interrater agreement and reliability. *Measurement in Physical Education and Exercise Science, 5*(1), 13-34.

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

Grunbaum, J. A., Kann, L., Kinchen, S., Williams, B., Ross, J. G., Lowry, R., et al.(2002). Youth risk behavior surveillance-United States, 2001. *Morbidity and Mortality Weekly Report, 51*(SS-4), 1-64.

Hatano, Y. (1993). Use of the pedometer for promoting daily walking exercise. *ICHPER*, *29*, 4-8.

Horvat, M., & Franklin, C. (2001). The effects of the environment on physical activity patterns of children with mental retardation. *Research Quarterly for Exercise and Sport, 72*(2), 185-195.

Janz, K.F., Dawson, J. D., & Mahoney, L. T. (2000). Tracking physical fitness and physical activity from childhood to adolescence: The Muscatine study. *Medicine and Science in Sports and Exercise*, *32*(7), 1250-1257.

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

Keating, X. D., Kulinna, P. H., & Silverman, S. (1999). Measuring teaching behaviors, lesson context, and physical activity in school physical education programs: comparing the SOFIT and the C-SOFIT instruments. *Measurement in Physical Education and Exercise Science, 3*(4), 207-220.

Kilanowski, C. K., Consalvi, A. R., & Epstein, L. H. (1999). Validation of an electronic pedometer for measurement of physical activity in children. *Pediatric Exercise Science, 11*, 63-68.

Kochersberger, G., McConnell, E., Kuchibhatla, M. N., & Pieper, C. (1996). The reliability, validity, and stability of a measure of physical activity in the elderly. *Archives of Physical Medicine and Rehabilitation, 77*, 793-795.

Kozub, F. M (2003). Explaining physical activity in individuals with mental retardation: An exploratory study. *Education and Training in Developmental Disabilities, 38*(3), 302-313.

Lakka, T. A., & Salonen, J. T. (1992). Intra-person variability of various physical activity assessments in the Kuopio ischemic heart disease risk factor study. *International Journal of Epidemiology*, *21*(3), 467-472.

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

Levinson, L., & Reid, G. (1991). Patterns of physical activity among youngsters with developmental disabilities. *Canadian Association for Health, Physical Education, and Recreation*, *56*, 24-28.

Lopez-Alarcon, M., Merrifield, J., Fields, D. A., Hilario-Hailey, T., Franklin, F., A., Shewchuk, R. M. et al. (2004). Ability of the Actiwatch accelerometer to predict free-living energy expenditure in young children. *Obesity Research, 12*(11), 1859-1865.

Lorenzi, D. G., Horvat, M., & Pellegrini, A. D. (2000). Physical activity of children with and without mental retardation in inclusive recess settings. *Educational Training in Mental Retardation and Developmental Disabilities, 35*(2), 160-167.

Louie, L., Eston R.G., Rowlands, A.V., Tong, K. K., Ingledew, D. K., & Fu, F. H. (1999). Validity of heart rate, pedometry, and accelerometry for estimating the energy cost of activity in Hong Kong Chinese boys. *Pediatric Exercise Science, 11*, 229-239.

Manions, Y. Kafatos, A., & Markakis, G. (1998). Physical activity of 6-year-old children: Validation of two proxy reports. *Pediatric Exercise Science, 10*(2), 176-188.

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.

McKenzie, T. L., Sallis, J. F., Nader, P. R., Patterson, T. L., Elder, J. P., Berry, C. C., et al. (1991). BEACHES: An observational system for assessing children's eating and physical activity behaviors and associated events. *Journal of Applied Behavior Analysis, 24*, 141-151.

Metcalf, B. S., Curnow, J. S., Evans, C., Voss, L. D., & Wilkin, T. J. (2002). Technical reliability of the CSA activity monitor: The EarlyBird Study. *Medicine and Science in Sports and Exercise, 34,* 1533-1537.

Minimitter Company Inc. (2000). *Actiwatch activity monitors: Instruction manual.* Sunriver, OR: Author.

Montoye, H. J. & Taylor, H. L. (1984). Measurement of physical activity in population studies: a review. *Human Biology*, *5*, 195-216.

Morrow, J. R. (1989). Generalizability theory. In M. J. Safrit & T. M. Wood (Eds.), *Measurement concepts in physical education and exercise science (pp.73-96).* Champaign, IL: Human Kinetics.

National Center for Health Statistics (1994). Plan and operation of the third National Health and Nutrition Examination Survey, 1998-94. *Vital Health Statistics, 1*, 39-41.

Nichols, J. F., Morgan, C. G., Sarkin, J. A., Sallis, J. F., & Calfas, K. J. (1999). Validity, reliability, and calibration of the Tritrac accelerometer as a measure of physical activity. *Medicine and Science in Sports and Exercise, 31*, 908-912.

Noland, M., Danner F., DeWalt K., McFadden M., & Kotchen J. M. (1990). The measurement of physical activity in young children. *Research Quarterly and Exercise Sport, 61*(2), 146-53.

Pate, R., Corbin, C., Pangrazi, B. (1998). Physical activity for young people. *President's Council on Physical Fitness and Sports Research Digest.* Retrieved August 26, 2005, from http://www.fitness.gov/digest_sep1998.htm .

Pate, P. R, Pratt M., Blair S. N., Haskell, W. L., Macera, C. A., & Bouchard, C. et al. (1995). A recommendation from the centers for disease control and prevention and the American college of sports medicine. *Journal of the American Medical Association, 273*(5), 402-407.

Puhl, J., Greaves, K., Baranowski, T., Gruben, D., & Seale, D. (1990). Descriptions and calibration of a Children's Activity Rating Scale (CARS). *Research Quarterly Exercise and Sports, 61,* 459-477.

Puyau, M. R., Adolph, A. L., Vohra, F. A., & Butte, N. F. (2002). Validation and calibration of physical activity monitors in children. *Obesity Research, 10*(3), 150-157.

Robertson, J., Emerson, E., Gregor, N., Hatton, C., Turner, S., Kessissoglou S., & Hallam, A. (2000). Lifestyle related risk factors for poor health in residential settings for people with intellectual disabilities. *Research in Developmental Disabilities, 21*, 469-486.

Rosser Sandt, D. D., & Frey, G. C. (2005). Comparison of physical activity levels between children with and without autistic spectrum disorders. *Adapted Physical Activity Quarterly, 22*, 146-159.

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

Sallis, J. F., Paterson, T. L., Buono, M. J., & Nader P. R. (1988). Relationship of cardiovascular fitness and physical activity to cardiovascular disease risk factors in children and adults. *American Journal of Epidemiology, 127*, 933-941.

Saris, W. H. M., Blair, S. N., & van Baak, M. A. et al. (2003). How much physical activity is enough to prevent unhealthy weight gain? Outcome of the IASO 1st Stock Conference and consensus statement. *Obesity Review, 4*, 101-114.

Schpherd, E. F., Toloza, E., McClung, C. D., & Schmalzried, T. P. (1999). Step activity monitor: Increased accuracy 8in quantifying ambulatory activity, *Journal of Orthopaedic Research, 17*, 703-708.

Scruggs, P. W., Beveridge, S. K., Eisenman, P. A., Watson, D. L., Shultz, B. B., & Ransdell, L. B. (2003). Quantifying physical activity via pedometry in elementary physical education. *Medicine and Science in Sport and Exercise*, *35*, 1065-1071

Shavelson, R. J. & Webb, N. M. (1991). Generalizability *Theory: A primer*. Newbury Park, CA: Sage.

Shavelson, R. J., Webb, N. M., & Rowley, G. L. (1989). Generalizability theory. *American Psychologist, 44*(6), 922-932.

Sieminki, D. J., Cowell, L. L., Montgomery, P. S., Pillai, S. B., & Gardner, A. W. (1997). Physical activity monitoring in patients with peripheral arterial occlusive disease. *Journal of Cardiopulmonary Rehabilitation, 17*, 43-47.

Sirard, J. R., & Pate, R. R. (2001). Physical activity assessment in children and adolescents. *Sports Medicine, 31*(6), 439-454.

Stanish, H. I. (2004). Accuracy of pedometers and walking activity in adults with mental retardation. *Adapted Physical Activity Quarterly, 21*, 167-179.

Stanish, H. I. & Draheim, C. C. (2005). Assessment of walking activity using a pedometer and survey in adults with mental retardation. *Adapted Physical Activity Quarterly, 22*, 136-145.

Suzuki, M., Saitoh, S., Tasaki, Y., Shimonmura, Y., Makishima, R., & Hosoya, N. (1991). Nutrition status and daily physical activity of handicapped students in Tokyo metropolitan schools for deaf, blind, mentally retarded, and physically handicapped individuals. *American Journal of Clinical Nutrition, 54*, 1101-1111.

Temple, V. A., Anderson, C., & Walkley, J. W. (2000). Physical activity levels of individuals living in a group home. *Journal of Intellectual and Developmental Disability, 25*, 327-341.

Temple, V. A. & Walkley, J. W. (2003). Physical activity of adults with intellectual disability. *Journal of Intellectual & Developmental Disability, 28*(4), 232-334.

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

Trost, S. G., Pate, P. R., Sallis, J. F., Freedson, P. S., Taylor, W. C., Dowda, M., et al. (2002). Age and gender differences in objectively measured physical activity in youth. *Medicine and Science in Sports and Exercise*, *34*(2), 350-355.

Trost, S. G., Ward, D. S., Moorehead, S. M., Watson, P. D., Riner, W., & Burke, J. R (1998). Validity of the Computer Science and Applications (CSA) activity monitor in children. *Medicine and Science in Sports and Exercise, 30*, 629-633.

Treuth, M. S. (2002). Applying multiple methods to improve the accuracy of activity assessments. In G. J. Welk (Ed.), *Physical activity assessments for health-related research (pp.213-225)*. Champaign, IL: Human Kinetics.

Tudor-Locke, C. E. (2002). Taking steps toward increased physical activity: Using pedometers to measure and motivate. *Council on Physical Fitness and Sports Research Digest*, *3*(17), 1-8.

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

Tudor-Locke, C. E., & Myers, A. M. (2001). Methodological considerations for researchers and practitioners using pedometers to measure physical (ambulatory) activity. *Research Quarterly for Exercise and Sport, 72*, 1-12.

Ulrich, D. A., & Wise, S. L. (1984). Reliability of scores obtained with the objectives based motor skill assessment instrument. *Adapted Physical Activity Quarterly, 1*, 230-239.

U.S. Department of Health and Human Services (1996). *Physical activity and health: A report of the Surgeon General.* Atlanta, GA: U.S. Department of HHS.

U.S. Department of Health and Human Services (2000). *Healthy people 2010.* Washington, DC: U.S. Gov. Printing Office.

Vincent, S. D., & Pangrazi, R. P. (2002). Does reactivity exit in children when measuring activity levels with pedometers? *Pediatric Exercise Science, 14,* 56-63.

Vincent, S. D., & Pangrazi, R. P., Raustorp, A., Tomson, L. M., & Cuddihy, T. F. (2003). Activity levels and body mass index of children in the United states, Sweden, and Australia. *Medicine and Science in Sports and Exercise, 35*(8), 1367 - 1373

Vincent, S. D. & Sidman, C. L. (2003). Determining measurement error in digital pedometers. *Measurement in Physical Education and Exercise Science, 7*, 19-24.

Welk, G. J., Schaben, J. A., & Morrow, J. R. (2004). Reliability of accelerometry-based activity monitors: A generalizabilty study. *Medicine and Science in Sports and Exercise, 36*(9), 1637-1645.

Welk, G. J., & Wood, K. (2000). Physical activity assessments in physical education: A practical review of instruments and their use in the curriculum. *Journal of Physical Education, Recreation, and Dance, 71*(1), 30-40.

APPENDICES

APPENDIX A - REVIEW OF THE LITERATURE

The purpose of this literature review is to provide an overview of published work in the areas of health promotion through physical activity for persons with developmental disabilities, measurements of physical activity, and the generalizability theory. This material provides the background information for the current study.

Physical Activity and Health

Recently, public health policy has shifted from improving health-related fitness to increasing regular physical activity for optimal health (Welk & Wood, 2000). Caspersen, Powell and Christen (1985) originally defined physical activity, physical fitness, and health-related fitness. According to them, physical activity is defined as "any bodily movement produced by skeletal muscle that results in energy expenditure" (p.126) while physical fitness is defined as "a set of attributes that people have or achieve" (p.128). Health-related fitness is a subcategory of physical fitness and has five components (cardiorespiratory fitness, muscular strength and endurance, body composition, and flexibility) that have a relationship with good health (Corbin, Pangrazi, & Franks, 2000)

Emphasizing the importance of regular physical activity is based on the significant relationships between lack of physical activity and the risk factors for obesity, coronary heart disease, diabetes, and colon cancer (Blair, 1993; Pate et al., 1995; U.S. Department of Health and Human Service [USDHHS], 1996). The Surgeon General's report on physical activity and health concluded that moderate physical activity can substantially reduce the risk of developing or dying from hearth disease, diabetes, colon cancer, and high blood pressure (USDHHS, 1996). Healthy People 2010 (USDHHS, 2000) identified physical activity as one of the ten leading health indicators.

Physical Activity Recommendation

Different public agencies have developed several physical activity guidelines. Compared to traditional exercise-fitness guidelines, new physical activity guidelines are different in terms of emphasizing moderate levels of physical activity and an accumulation of physical activity in intermittent or short bouts. The Surgeon General's report on physical activity and health (USDHHS, 1996) developed physical activity guidelines for Americans over the age of two. It recommended that all Americans over the age of two need to accumulate 30 minutes of moderate activity on most days of the week; moderate levels of physical activity is "roughly equivalent to physical activity that uses approximately 150 kcal of energy per day, or 1,000 kcal per week" (USDHHS, 1996).

However, it has been proposed that the recommendation is not sufficient for children (Cavill, Biddle, & Sallis, 2001; Pate, Corbin & Pangrazi, 1998, Saris et al., 2003), or preventing unhealthy weight gain (Saris et al., 2003). Saris et al. (2003) proposed that adults, who were previously obese, need to participate in 60-90 minutes of MVPA to prevent weight regain, and participating 45 to 60 minutes of moderate levels of physical activity per day can prevent the transition to overweight or obesity. For children (5-18 years), recent guidelines recommend that youth need to participate in at least 60 minutes of physical activity daily (Cavill, Biddle, & Sallis, 2001; Pate, Corbin, & Pangrazi, 1998). In addition to the recommendation, the Council for Physical Education for Children (COPEC) provided detail physical activity guidelines for children from 5 to 12 years old; children should participate in multiple sessions of moderate to vigorous physical activity (MVPA) lasting 10-15 min per session (COPEC, 1998).

There is also pedometer-determined physical activity recommendation. Accumulating 10,000 steps per day has been promoted in Japan since 1960s, and gained popularity with the media and in practice in the U.S. (Tudor-Locke, & Bassett Jr., 2004). The value of 10,000 steps per day has been proposed to an appropriate goal for everyone (Hatano, 1993). The assumption of the value is based on that it is approximately equivalent of an energy expenditure of 336 (slow walking), 382 (fairly fast walking), and 431 kcal (fast walking) per day (Hatano, 1993).

However, it has been questioned to recommend the value of 10,000 steps per day to children (Tudor-Locke & Myers, 2001; Tudor-Locke, 2002) because children usually accumulate more than 10,000 steps per day. It has been reported that children (6-12 years) girls and boys take 10,661-11,383 and 12,554-13,871 steps per day, respectively (Vincent, Pangrazi, Raustorp, Tomson, & Cuddihy, 2003).

Measurement Tools of Physical Activity

A variety of methods (self-report, activity monitors, heart rate monitor, pedometers, direct observation, indirect calorimetry, and doubly labeled water) have been used to measure physical activity for children. Each method has advantages and disadvantages. Accelerometers, pedometers, and proxy reports have been commonly employed to measure physical activity levels of individuals with and without DD.

Accelerometers

Accelerometers use an electronic component to assess the acceleration of the body movement either uniaxial or in multiple dimensions (Dale, Welk, & Matthews, 2002). These devices have been developed to measure frequency and intensity of body movement. The data of body movement can be stored over time and exported using computer software programs. However, these devices have a number of disadvantages; they are not suitable for aquatic activities, and have difficulty measuring static activities or activities were there is minimal movement of the body's center of gravity, such as rowing or cycling (Dishman, Washburn, & Scholeller, 2001).

The Actiwatch accelerometer (Minimitter, Bend, OR) has an omnidirectional sensor that is able to detect acceleration in all directions, but most sensitive in the direction parallel with the longest dimension of the case (Puyau, Adolph, Vohra, & Butte, 2002). An increased degree of speed and motion produces an increase in voltage, and the Actiwatch monitor stores this information as activity counts (Lopez-Alarcon et al., 2004).

Mixed results have been found on accuracy of Actiwatch in terms of estimating energy expenditure in children. Lopez-Alarcon et al. (2004) compared activity counts of Actiwatch (worn on the right ankle over 8 days) to total energy expenditure (TEE) measured with the doubly labeled water technique in children (from 4 to 6 years old).

There were no significant relationships between the averages of total daily activity counts and TEE. However, Puyau et al. (2002) found high correlations between activity counts (worn on hip and leg) and 6-hour energy expenditure (measured by room respiration calorimetry) in children (ages 6 to 16 years); Actiwatch on hip (r = .78), and leg (r = .80). Also high correlations were found between activity counts of Actiwatch and the Computer Science and Applications Actigraph (CSA) monitor on hip (r = .88), and on the leg (r = .89).

Pedometer

Pedometers are relatively simple, inexpensive and objective electronic devices measuring ambulatory behaviors, and record body movement as steps taken (Sirard & Pate, 2001). Pedometers have a horizontal, spring suspended lever arm that moves up and down with normal ambulation, and an accumulated step is displayed digitally on the screen of pedometers (Tudor-Locke, 2002). However, most pedometers can detect only total steps over the observation period, and cannot assess the intensity or pattern of physical activities, and are not sensitive to non-ambulatory activities (Sirard & Pate, 2001; Tudor-Locke, 2002). The Computer Science and Applications (CSA) actimeter (CSA model 7164) has the capacity to store the number of steps per minute over a 21-day period, which allow physical activity patterns to be examined , but it costs much more than regular pedometers ($325 for the CSA actimeter and $450 for the computer interface) (Bassett & Strath, 2002).

Validity evidence of pedometers in measuring physical activity levels of adults with MR and children without disabilities have been examined by comparing pedometer step counts with hand-tallied step counts. Stanish (2004) found high correlation (*r* > .95) between step counts and hand-tallied step counts during normal paced and fast paced walks on indoor and outdoor tracks. Moreover, wearing pedometers on either the right or left side did not have an effect on the accuracy of the pedometers. There has been no research examining validity evidences of pedometers for children with DD, but there has been for children without disabilities. Beets, Patton, and Edwards (2005) examined the effect of walking speed on the accuracy of pedometers, comparing it to hand-tallied step counts. In the study, twenty children (5-11 years old) were asked to walk "at their normal pace" an outdoor athletic track (400m) and on a treadmill at various speed (i.e., 40, 54, 67, 80, and 94 m/min). Two brands of pedometers (Walk4Life 2505, and Digiwalker SW-200) were used during the self-paced walking (SPW), and 4 models (Walk4Life 2505, Digiwalker SW-200, Sun TrekLINQ and Digiwalker SW-701) were used for treadmill walking. High intraclass correlation coefficients (ICC) (*R* =.99) were found between actual steps and pedometer step counts for both models across 3 SPW trials. In treadmill walking, high ICCs (*R* > .90) were examined except one model (SUN) at speed 67, 80, and 94 m/min. However, low ICCs (*R* < .72) were observed for all models at speed 54 and 40 m/min indicating less accuracy of pedometers during slow waking.

Relationships between the outputs of pedometers and accelerometers are different based on intensity levels of physical activities that children perform. The results of study by Kilanowski, Consalvi, and Epstein (1999) revealed significant and high correlations between pedometers (Digiwalker SW-200) and tri-axial accelerometers during combined classroom and recreation settings (*r* = .99), and recreational period alone (*r* = .98). However, low correlation was found in the classroom setting (*r* = .50) indicating less sensitivity of pedometers during low intensity activities.

Significant correlation has been found between pedometers and systematic observation. Kilanowski et al. (1999) compared pedometer steps with the results of systematic observations using the Children's Activity Rating Scale (CARS) (Puhl, Greaves, Baranowski, Gruben, & Seale, 1990). High correlations were found in combined classroom and recreation (*r* = .96), recreational setting (*r* = .97), and classroom (*r* = .80) settings. Using the Computerized System for Observing Fitness Instruction Time (C-SOFIT) (Keating, Kulinna, & Silverman, 1999), Scruggs et al. (2003) also found moderately high correlations (*r* = .74 - .85) between pedometer steps per minute and percent MVPA for children in first and second grade in physical education classes.

Similar results have been reported in relationships between pedometer and measures of energy expenditure. Correlations of pedometers (worn on left hip) with oxygen uptake scaled for body mass were examined for Welsh (Eston, Rowlands, & Ingledew, 1998) and for Hong Kong Chinese (Louie et al., 1999) children using identical measurement tools and procedures. Significant correlations (Welsh vs. Hong Kong children) were reported during treadmill activities (i.e. walking and running) (*r* = .78 vs. .77), unregulated play activities (i.e. catching, playing hopscotch, and sitting and coloring) (*r* = .921 vs. .931), and all activities combined (*r* = .81 vs. .86). Moreover, the results of simple liner regression equations to predict energy expenditure indicated that pedometers accounted for 73% and 65% (Hong Kong vs. Welsh children) of the variance.

Proxy Reports

It has been suggested that teacher and parent reports on children's physical activity levels can be used as useful tools providing physical activity levels of younger children (Sallis, Paterson, Buono, & Nader, 1988). The main advantage of proxy report is that recall errors due to individual's limited cognition can be decreased (Sirard & Pate, 2001). However, very limited information related to validity and reliability evidence of proxy reports on physical activity levels for individuals with DD is available. Temple and Walkely (2003) reported significant correlation (R = .78) between two energy expenditure estimates generated with a proxy report, the Bouchard Three-Day Physical Activity Record (Bouchard, Tremblay, Leblanc, Lortie, & Savard, 1983), and accelerometers.

A few studies established validity evidence of proxy reports for children without disabilities. Manions, Kafatos and Markakis (1998) compared the results of two proxy reports to heart rate data for 6-year-old children in Greece. Teachers reported MVPA levels of the children over 5 days based on the physical activity of the children during school while parents/guardians reported MVPA levels of the children during 3 days (2 consecutive weekdays, and one day of the weekend) based on the physical activity of the children after school. Based on the teacher's report, the children were categorized into five groups. Significant correlations were found between heart rate data (average minutes of MVPA over 3 days) and teachers' classification of the children during the 3 day (r = .41), during 2 school days (r = .59), and during 2 weekdays (r = .45). However, an insignificant relationship was found during one weekend day. Significant and higher correlations were found between heart rate data and parent reports (the numbers of reported 30-min sessions of MVPA during the after school period) during the first week day of after school period (r = .72), and during the second week day of after school period (r = .76), and during one weekend (r = .82). Noland, Danner, DeWalt, McFadden, and Kotchen (1990) examined whether physical activity levels of children (from 3 to 5 years old) reported by their parents and teachers can predict observed physical activity levels (measured with systematic observation) of the children during home (for 6 hours) and a playroom (for 20 minutes), and found no relationships between the two proxy reports and observed physical activity levels of children.

Issues in Physical Activity Measurement

Intra-Individual Variability

Intra-individual variability (IIV) in physical activity behaviors is problematic for physical activity research (Dale, Welk, & Matthews, 2002). IIV in physical activity research indicates difference of physical activity behaviors in a person from day to day (Baranowski & Moor, 2000). In order to obtain more representative data of daily physical activity patterns, physical activity must be monitored for more than 1 day (Bassett & Strath, 2002). To estimate consistency of physical activity behaviors across days, intra-class correlation coefficients (R) have been used. According to Baranowski and de Moor (2000), a multiple day intra-class correlation coefficient of .80 is the usual desired value.

No studies have examined the minimal number of monitoring days for assessing daily physical activity levels of individuals with DD to get reliable information of their habitual or usual physical activity patterns, but a few studies examined it for individuals without disabilities. Although it has been suggested to use multiple measures to reduce the error in measurement of daily physical activity patterns, there is no research that has reported IIV for combined physical activity measurement tools.

Different recommendations have been proposed based on a population and instruments that were used for the studies. Using pedometers, 6 days (R =.83) for middle school students (Rowe, Mahar, Raedeke, & Lore, 2004) and 5 days (R =.82) for elementary students (Vincent & Pangrazi, 2002) were recommended. Using uniaxial accelerometer, Janz, Witt, and Mahoney (1995) recommended monitoring daily physical activity levels of children (between 7 and 12 years old) for 4 days if reliability of .70 is acceptable. To achieve reliability of .80, six days of monitoring was recommended. Also, Trost, Pate, Freedson, Sallis, and Taylor (2000) examined the required number of days for monitoring daily physical activity using Spearman-Brown prophecy formula (Stanley, 1971). The results of the Spearman Brown prophecy formula indicated that 3 days and 5 days of monitoring were required to get reliability of .70 and .80 for children in grade one through six. However, children in grades seven through twelve were required to be monitored for 5 days and 9 days in order to achieve reliability of .70 and .80.

One study examined the minimum days of monitoring daily physical activity levels of adults. Gretebeck and Montoye (1992) monitored physical activity of adults without disabilities (24 - 67 years) for 7 consecutive days using pedometers and triaxial accelerometers (Caltrac). Significantly less physical activity during weekends than weekdays were examined. The results of Spearman-Brown Prophecy formula indicated that at least 4 days and 5 days are needed to minimize the IIV using pedometers worn on ankle and waist, respectively, while 2 days are required using Caltrac (wearing on ankle).

Lack of Criterion Measurement

Due to lack of criterion physical activity measurement tools, multiple methods have been recommended to increase the accuracy of physical activity assessments (Treuth, 2002). A few studies examined the accuracy of physical activity using combined methods. Most of the studies evaluated the use of heart rate monitors with other activity sensors to estimate energy expenditure more accurately. Eston et al. (1998) computed multiple regression equations to predict oxygen uptake from pairs of measurement among pedometers (Digiwalker DW-200), uniaxial accelerometers (WAM, model 7164), Tritrac-R3D accelerometers (model T303), and heart rate monitors (BHL 6000 Medical). Oxygen uptake was measured while Welsh children (N = 30, aged 8.2 - 10.8) were walking at two speed (4 and 6 km/h), running at two speeds (8 and 10 km/h) on a treadmill, playing catch, hopscotch, and sitting and crayoning. Children wore 3 pedometers, one on their left ankle, one on their wrist, and one on their waist. Both accelerometers were set to collect data every 1 minute and were worn above the left hip of each child. Louie et al. (1999) duplicated the study by Eston et al. (1999) for 21 Hong Kong Chinese boys, aged 8-10 years. Both studies found similar results. The best combination among measurement tools, the Tritrac-R3D and heart rate monitor was the strongest combined predictors (R^2 = .85 for Welsh children, R^2 = .92 for Chinese children). For the pedometer on hip and uniaxial accelerometer (R^2 = .78 for Welsh children, R^2 = .84 for Chinese children) in which the hip pedometer added 16.6% (Welsh children) and 10.6% (Chinese children) more to the variance than the uniaxial accelerometer alone. In contrast, the variance increased by 12.2% and 10.8% when the uniaxial accelerometer was added to the hip pedometer.

Using two different types of accelerometers including a triaxial accelerometer (Tritrac-R3D) and a wrist-worn uniaxial accelerometer (ActiWatch), Chen et al. (2003) investigated whether the prediction of energy expenditure of physical activity for 60 sedentary females without disabilities would be improved compared to the use of only one

accelerometer. Energy expenditure during a 24-h period was measured using a whole-room indirect calorimeter. The results of this study indicated that the combination of the two accelerometers predict total energy expenditure (97.7 %) significantly better than using either Tritra-R3D (90%) or ActiWatch (86%).

Generalizability Theory

Generalizability theory (G-theory), which is an extension of classical test theory, allows researchers to estimate reliability while identifying multiple sources of error (facet) separately in a single model (Goodwin, 2001; Morrow, 1989; Shavelson, Webb, & Rowley, 1989). According to Shavelson, Webb, and Rowley (1989), "Instead of asking how accurately observed scores reflect their corresponding true scores, G-theory asks how accurately observed scores permit us to generalize about persons' behavior in a defined universe of situation." (p.92). GT provides a "generalizability (G) coefficient" (from zero to one), and the magnitude of each source of error (Shavelson, Webb, & Rowley, 1989).

GT distinguishes between generalizability (G) and decision (D) studies; G study is designed to estimate the magnitude of as many potential sources of error as possible, while D studies are used to make decisions based on information from a G study to design optimal measurement protocol for a particular purpose (Shavelson & Webb, 1991). G studies can be conducted using crossed, nested or mixed models; in crossed models, all of facets are regarded as random while in nested models some facets are fixed within other facets and in mixed models, some facets are random and some are fixed (Morrow, 1989).

This measurement approach is very useful in the assessment of physical activity because various sources of variability (i.e. day, instrument etc.) in habitual physical activity behavior can be identified, and thereby can reduce the influence of these sources on the measurement. However, the theory remains relatively unused. Welk, Schaben and Morrow (2004) conducted a G-study to examine variability in accelerometer counts; Two-facets (4 monitors, 3 trials) in a fully crossed design were employed. College students were asked to walk on a treadmill (3 mph) 3 times while wearing 4 different types of accerelerometers (Tritrac, CSA/MTI, Biotrainer, and Actical), 1% (CSA/MTI), 9% (Biotrainer), 10% (Actical), and 12% (Tritrac) of the variance were due to monitors. Relatively large amount of variance occurs due to the interaction between subjects and trials (from 14% to 21%), and among monitors, trials and subjects (from 18% to 23%). The results of this study suggest that correct and consistent positioning of accelerometers across trials is important to decrease measurement error.

Physical Activity of Individuals with Developmental Disorders

Adults

Limited information is available regarding daily physical activity levels of adults with developmental disabilities (DD). A majority of adults with mental retardation (MR) do not meet the current recommendation of physical activity of participating 30 min of MVPA for five or more days per week (Temple, Anderson, & Walkley, 2000; Temple & Walkley, 2003) while mixed results have been reported regarding the accumulation of 10,000 steps per day (Stanish, 2004; Stanish & Draheim, 2005). Stanish (2004) examined daily steps of adults with mild MR (19-65 years) (including Down syndrome) over 7 days. They reported that the average number of steps/day was 10,811 and 7,863 for females and males, respectively. Forty-five % of the participants accumulated over 10,000 steps per day during weekdays, but 20% of them did over 10,000 steps per day during weekends. Also, the participants with Down Syndrome (DS) walked significantly less than those with MR; the average number of steps/day for 1) males with and without DS were 11, 885 and 5,449, respectively, and 2) females without and with DS were 11,809 and 8,815, respectively. In another study, adults with mild and moderate MR (between 19 and 65 years of age) wore a pedometer for 7 consecutive days, and the average of steps per day was 7,832 steps (Stanish & Draheim, 2005). Only 21% of the participants accumulated more than 10,000 steps/day.

Some studies examined whether adults with MR meet the recommendation of participating in MVPA for 30 minutes on most of days of week using accelerometers. Most of the adults with MR engaged in insufficient regular physical activity. Temple and Walkley (2003) examined physical activity of Australians with MR in supported group homes for 3 consecutive days. Only 32% of the participants met the recommendation. Similar results were reported by Temple, Anderson, and Walkley (2000) who monitored physical activity of Australian individuals with MR (19-45 years) in a group home for 7 consecutive days. Only two participants out of six (about 33%) met the recommendation. Interestingly, three participants reported that they walked an average of 107, 98, and 84 min/day, but the level of intensity was not sufficient enough to consider MVPA.

Other studies examined the prevalence of physical inactivity of adults with MR using questionnaires which were developed for adults without disabilities. Using, Health Surveys for England 1993 (Department of Health, 1995), Robertson et al. (2000) examined physical activity levels of adults with MR, who lived in different environments including village communities, residential campuses, and dispersed housing schemes. The participants in residential campuses were the most inactive group, while those in dispersed housing were less inactive than people in village communities; 88%, 93%, and 80% of participants (village communities, residential campuses, and dispersed housing schemes, respectively) engaged in MVPA (over 5 kcals/min and lasted for 20 min or more) less than 12 times, which was consider inactive by the survey. Using the National Health and Nutrition Examination Survey III and the Physical Activity Survey (NHANES III; National Center for Health Statistics, 1994), Draheim, Williams, and McCubbin (2002) examined leisure time physical activity (LTPA) of adults with MR (19-65 years). The results of the surveys revealed that 13% and 49% of the participants engaged in no LTPA and little to no LTPA (participating in moderate to vigorous LTPA less than 3 times/week), respectively.

Children

There is the limited available information regarding physical activity patterns of children with DD. A few studies have examined physical activity patterns of children with DD throughout the day (Kozub, 2003; Rosser Sandt & Frey 2005), in school settings (Horvat & Franklin, 2001), and during their leisure time (Levinson & Reid, 1991). Kozub (2003) reported that adolescents with MR were mostly active during the afternoon, and engaged in moderate levels of physical activity lasting short periods of time (from 2 to 4 min), and younger children more engaged in moderate levels of physical activity compared to older children. The daily physical activity levels of the adolescents with MR were measured with the TriTrac R3D monitors over seven days. Rosser Sandt and Frey (2005) found that children with autistic spectrum disorders (ASD) ages 5 to 12 years engaged in at least 60 minutes on MVPA during the five days.

Their physical activity levels were measured with uni-axial accelerometers and direct observation using the Behavior of Eating and Activity for Children's Health: Evaluation System (BEACHES; McKenzie et al., 1991). Moreover, the children participated in significantly more percent time in MVPA during recess in school than after school.

In a school setting, children with DD spend significantly more percent time in MVPA during recess than physical education. Horvat and Franklin (2001) found that physical activity levels of children with MR (grade kindergarten through five) were significantly higher in recess settings than classroom settings based on the data of heart rate monitors, accelerometers and the Scheme for Observing Activity Level (SOAL) checklist (Eaton, Enns, & Presse, 1987). Also, Rosser Sandt and Frey (2005) reported that children with ASD spent significantly more percent time in MVPA during recess than physical education.

Limited information is available regarding the physical activity patterns of children with DD during their leisure time. Levinson and Reid (1991) examined physical activity leisure patterns of youngsters with DD (from 4 to 21 years old) including MR, emotional disturbance, varying degrees of autism, and/or mild neurological impairments via parent reports. Seventy five % of parents of younger students (4-10 years) placed their children in the activity category compared to 56% for older students (11-21 years). The five most common activities of the participants were walking, swimming, bicycling, jogging, running, and ice-skating. Moreover, the younger students were active at home (95%), at the park (84%), and at school (74%).

Older students were active in community facilities (67%), school (63%) and at home (63%).

Whether accumulating 10,000 steps per day is a reasonable goal for children with DD should be examined. It may be low for children with DD because walking is the most common physical activity for them (Levinson & Reid, 1991). Suzuki et al. (1991) monitored ambulatory physical activity of children with MR (from 3 to 22 years old) over 6 days using pedometers, and found that they walk over 10,000 steps per day. Males with MR (*M* = 16,000 steps per day) walked significantly more than females with MR (*M* = 12,300 steps per day). Significantly negative correlations were found between steps and age for each gender.

Mixed results have been found related to a disparity of physical activity levels between children with and without DD during school and throughout the day. No significant difference of physical activity levels during school was found by Faison-Hodge and Porretta (2004) as they measured the physical activity levels of children with and without mental retardation (MR), grades third through fifth, using the System for Observing Fitness Instruction Time (SOFIT) (McKenzie, Sallis, & Nader, 1991). Physical activity levels of children with MR during physical education and recess settings were compared with non disabled children with low and high cardiorespiratory function. No significant differences were found among the three groups of children in both settings. Also, Lorenzi et al. (2000) found no group differences in the physical activity levels of children with and without MR (grades kindergarten through five) during recess based on the results of the observation using the Scheme for Observing Activity Level (SOAL) checklist (Eaton, Enns, & Presse, 1987). However, the data from heart rate monitors and accelerometers demonstrated that male students with MR had higher levels of physical activity compared to male students without MR during recess.

There were no significant differences in physical activity levels between the children with and without ASD found (Rosser Sandt & Frey, 2005). Their physical activity levels were measured with uni-axial accelerometers and direct observation across four time periods including all day, after school, physical education class, and recess. Significant differences of daily physical activity levels of children with and without Prader-Willi Syndrome (PWS) were found. Eiholzer et al. (2003) found that children (aged 4-19 years) with PWS walked less (11.1 km vs. 24.6 km) ($p < .05$), and participated in lower levels of physical activity (256 points vs. 274 points) ($p < .01$) than those without disabilities. The participants' daily physical activity levels were examined with a pedometer over 3 days. Also, the children's parents or children supervised by their parents were required to rate their physical activity on the activity protocols using four scales from 1 to 4 in each half-hour of the day

APPENDIX B - INSTITUTIONAL REVIEW BORAD APPROVAL

OSU Oregon State UNIVERSITY

Institutional Review Board • Office of Sponsored Programs and Research Compliance
Oregon State University, 312 Kerr Administration Building, Corvallis, Oregon 97331-2140
Tel 541-737-8008 | Fax 541-737-3093 | http://oregonstate.edu/research/osprc/rc/humansubjects.htm
IRB@oregonstate.edu

TO: Joonkoo Yun,
Nutrition & Exercise Sciences

RE: Source of Variability in Daily Physical Activity for Adolescents with and without Developmental Disabilities
(Student Researcher: So-Yeun Kim)

IRB Application No. 3161

Level of Review: Expedited

Expiration Date: **2/21/2007**

The referenced project was reviewed under the guidelines of Oregon State University's Institutional Review Board (IRB). The IRB has **approved** the **new** request. A copy of this information will be provided to the full IRB committee.

Enclosed is the informed consent information for this project, which has received the IRB stamp. All participants must receive the IRB-stamped informed consent document.

- **MODIFICATION REQUEST:** Any proposed changes to the approved protocol, informed consent form(s), testing instrument(s), research staff, or increase in the number of participants must be submitted for approval before implementation.
- **ADVERSE EVENT:** Adverse Events must be reported within three days of occurrence. This includes any outcome that is not expected and routine and that results in bodily injury and/or psychological, emotional, or physical harm or stress.
- **CONTINUING REVIEW:** Before the expiration date noted above, a notice will be sent to remind researchers to complete the continuing review form to renew this project. It is imperative that the Continuing Review is completed and submitted by the due date in the notice or approval will lapse, resulting in a suspension of all activity.

Forms and additional information can be found at the IRB web site at: http://oregonstate.edu/research/osprc/rc/humansubjects.htm.

If you have any questions, please contact the IRB Human Protections Administrator at IRB@oregonstate.edu or by phone at (541) 737-8008.

Elisa Espinoza Fallows   Date: 2/22/06

Elisa Espinoza Fallows
IRB Human Protections Administrator

pc: 3161 file

APPENDIX C - PHYSICAL ACTIVITY LOG

Physical Activity Log

Name: ______________________ DOB: __________ Grade: ________
Date: ________________ Weather: ______________
* ***Wake-up time***: ____________ ***Bed-Time***: _____________
* School day or Non school day (please circle it)

Please record the types of physical activities, and for how long and where you participated in the physical activities.

Example;

Time	Physical Activity	Duration	Place
7	Walking	20 min	From home to school
8			
9	Basketball	30 min	PE class, gym at school

Time	Physical Activity	Duration	Place
7			
8			
9			
10			
11			
Noon			
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			
11			
12AM			

APPENDIX D - REMINDER POSTER

Don't forget to wear

Pedometers

✪ <u>NO</u> Waterproof!

APPENDIX E – REMINDER CALLS

To participants without DD

First day:

"Good morning, this is So-Yeun Kim from Oregon State University. May I speak to (child's name). I am calling to remind you to wear pedometers and Actiwatches today. Thank you. Have a nice day."

Other days:

"Good morning, this is So-Yeun Kim from Oregon State University. May I speak to (child's name). I am calling to remind you to wear pedometers and Actiwatches today. Did you fill out the physical activity log?"

If the child says yes,

"Thank you. Have a nice day"

If the child says no,

"I would really appreciate it if you would fill out the physical activity log every day. Thank you. Have a nice day."

To participants with DD

First day:

"Good morning, this is So-Yeun Kim from Oregon State University. May I speak to (parent/guardian's name). I am calling to remind you to help your child to wear pedometers and Actiwatches today. Thank you. Have a nice day."

Other days:

"Good morning, this is So-Yeun Kim from Oregon State University. May I speak to (parent/guardian's name). I am calling to remind you to help your child to wear pedometers and Actiwatches today. Did you fill out the physical activity log last night?"

If the parent says yes,

"Thank you. Have a nice day"

If the parent says no,

""I would really appreciate it if you would fill out the physical activity log every day. Thank you. Have a nice day."

APPENDIX F – SAS COMMAND FILE FOR G-THEORY ANALYSIS

```
options ps = 55;
options ls = 79;
pageno = 1;
data pddw;
input
      sub i1d1 i1d2 i1d3 i1d4 i1d5 i2d1 i2d2 i2d3 i2d4 i2d5;
      instrument=1;day=1;score=i1d1;output;
      instrument=1;day=2;score=i1d2;output;
      instrument=1;day=3;score=i1d3;output;
      instrument=1;day=4;score=i1d4;output;
      instrument=1;day=5;score=i1d5;output;
      instrument=2;day=1;score=i2d1;output;
      instrument=2;day=2;score=i2d2;output;
      instrument=2;day=3;score=i2d3;output;
      instrument=2;day=4;score=i2d4;output;
      instrument=2;day=5;score=i2d5;output;

cards;
1 4418 5688 6674 9689 5680 4641 5875 6953 9468 5701
2 9782 7091 9119 9661 10236 9365 6758 9992 10268 10873
3 6090 7466 8375 6043 5135 5634 7267 7610 6724 5542
4 11171 10036 12999 8744 8478 10674 9783 12428 8806 8443
5 7519 5673 8104 5574 6561 7204 5565 8495 5407 6533
6 4346 11533 4610 7012 6020 4659 13181 3890 8411 6198
7 12175 9051 9275 10120 13315 12480 8212 9088 10546 12363
8 12131 10756 13721 12378 11047 11006 10457 14195 11769 9765
9 10464 7981 9408 8557 6663 10032 7274 9500 8674 6724
10 6666 6571 7123 5097 6140 6627 7799 7547 6243 6142
11 6021 5977 8212 9770 5049 5995 6344 7813 10078 4987
12 10082 11740 10100 13522 9619 9620 11396 10958 12320 10544
13 5940 5864 5028 4664 7450 5977 5232 5406 5466 6873
14 8521 13496 11106 11908 7430 8139 14319 11544 11165 7858
15 6428 8025 7225 6686 8520 5863 8067 6959 6688 8563
16 6955 7141 7693 7160 8209 6551 6841 8036 7062 8735
;

Proc varcomp method=type1;
      class sub instrument day;
      model score=sub instrument day
                  sub*instrument
                  sub*day
                  instrument*day;
Proc glm;
      class sub instrument day;
      model score=sub instrument day
                  sub*instrument
                  sub*day
                  instrument*day;
run;
```