FACTOR VALIDITY, INVARIANCE, AND COMPARISON OF SEVERAL

MEASURES OF PHYSICAL ACTIVITY ENJOYMENT

by

DEREK PATRICK HALES

(Under the direction of TED BAUMGARTNER)

ABSTRACT

Enjoyment is often cited as an important correlate of physical activity behavior and is included in many health promotion models and theories of behavior motivation. A number of measures (scales) have been developed and used to quantify enjoyment of physical activity but most lack sufficient validity evidence. In addition, little is known about how the various measures interrelate. Therefore, the purpose of this research study was to compare and extend the construct validity evidence for several measures of physical activity enjoyment. Five enjoyment scales were chosen for comparison. Confirmatory factor analysis was used to examine the factor validity, invariance, and convergence of the enjoyment scales. If the hypothesized factor structure failed to fit the data, exploratory analyses were used to determine a better fitting model. In addition, each measure of enjoyment was correlated with several estimates of physical activity, perceived competence and extrinsic motives for physical activity. Participants (N = 1023; mean age 19.60 ± 1.55 years) were recruited from basic physical education classes and the psychology research pool at two universities. A simple single factor model did not adequately account for the covariance among the items on four of the five scales. Correlated uniquenesses were added to the models for the PACES-M, GEQ, and IMI-joy to improve model fit. Based on exploratory analyses, a two

factor model was found to represent the data from the PACES. The factorial invariance

between black and white and between male and female participants was supported for all

scales. The results support the similarity of enjoyment scale scores, but suggest that

measuring enjoyment and perceived competence with the same instrument (i.e IMI and

MPAMR) inflates their relationship. Each measure of enjoyment was modestly correlated

with total and vigorous physical activity (0.22 – 0.34), but only slightly related to work and

moderate activity (0.047 – 0.12). As expected, the correlations of enjoyment with perceived

competence were larger than with the extrinsic motives. The magnitude of these

relationships was very similar across the five measures of enjoyment. Overall, this research

provides support for the construct validity of scores from several measures of physical

activity enjoyment, supporting similarity of measurement but not complete convergence.

FACTOR VALIDITY, INVARIANCE, AND COMPARISON OF SEVERAL

MEASURES OF PHYSICAL ACTIVITY ENJOYMENT

by

DEREK PATRICK HALES

B.A., University of North Carolina at Chapel Hill, 1996

M.A., East Carolina University, 2000

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial

Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2005

FACTOR VALIDITY, INVARIANCE, AND COMPARISON OF SEVERAL

MEASURES OF PHYSICAL ACTIVITY ENJOYMENT

by

DEREK PATRICK HALES

Major Professor:	Ted Baumgartner
Committee:	Deborah Bandalos
	Kirk J. Cureton
	Rod K. Dishman

Electronic Version Approved

Maureen Grasso
Dean of the Graduate School
The University of Georgia
August 2005

DEDICATION

For Shannon
You have been very patient. Your smile moves me. Your love inspires me.

For Lucas and Maggie
You may never read this, but know that you have my heart forever.

ACKNOWLEDGMENTS

Thanks to Dr. Baumgartner for bringing me here site-unseen. Your guidance has been very helpful and your willingness to help is a credit. Thanks for seeing me through the final blitz.

Dr. Dishman, thanks for giving me a job and acting as a mentor. I will always owe you a debit of gratitude. Guess what? Your job is not over.

Dr. Cureton, thanks for asking me to teach. Without it, I might have been done sooner, but I would not have had as much fun or gained as much experience. Also thanks for reading this thing.

Dr. Bandalos, you taught me most of what this whole thing is about, the method behind the madness. And like it or not, I'm hooked. A special thanks for working within my tight time frame.

To all my committee members, your service is greatly appreciated. I hope our paths continue to cross for years to come. I probably owe you all diner.

To my students, former and future, if you're reading this you've either been in school too long or you've got too much time on your hands. Go hear some music… no, no… go study.

To everyone that helped with my data collection, especially Steve Quinlan, Chris Henrich, and Kim Darnell, without your assistance this whole thing might not have gotten off the ground. May research karma bless you for years to come.

To everyone that completed my survey, I don't really know you, but I am intimately familiar with your enjoyment of physical activity.

To my family, thanks for all your love and support. You can finally stop saying "Uhh…He's still a student".

To Shannon, more than a man that's ten feet tall.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

Despite the well established benefits of a physically active lifestyle, Americans are generally inactive (US Department of Health and Human Services [USDHHS], 1996). Current estimates indicate that between 30% and 50% of adults get the recommended amount of physical activity, while 25% report no leisure time physical activity (USDHHS, 2003a; USDHHS, 2004). Alarmingly, low levels of physical activity are increasing among youth as well. It is estimated that 40% of girls and 27% of boys do not get the recommended amount of moderate or vigorous physical activity (USDHHS, 2003b). These low levels of physical activity, in conjunction with the consistent evidence that physically active individuals are healthier than their sedentary counterparts (USDHHS, 1996) has lead to the establishment of national physical activity goals and a multitude of interventions designed to increase physical activity levels. The risk of developing heart disease (Blair et al., 1989; Manson et al., 1999), hypertension (Paffenbarger, Wing, Hyde, & Jung, 1983) obesity (Williamson, 1993), diabetes (Helmrich, Ragland, Leung, & Paffenbarger, 1991; Knowler et al., 2002), and Depression (Dunn, Trivedi, & O'Neal, 2001) has consistently been shown to be lower for people who meet, or exceed, levels of physical activity currently suggested by the surgeon general. Despite believing that being physically active is important for health, it is obvious that most Americans are not motivated to accumulate enough physical activity to reap the benefits (Haase, Steptoe, Sallis, & Wardle, 2004).

For years, theories of motivation have been developed and used to explain the complexity of what moves people to act and the direction and persistence of their actions. Biological needs and

drives and psychological constructs such as competence and self-efficacy have been cited as key motivating factors for human behavior. Although debated, current theories generally categorize sources of motivation as extrinsic or intrinsic (Deci, 1975; Petri & Govern, 2004; Reiss, 2004). Although both can energize and direct behavior, intrinsic motivation is often considered the primary determinant of sustained human behavior (Deci & Ryan, 1985; Iso-Ahola & St. Clair, 2000; Ryan & Deci, 2000; Wankel, 1993).

One primary determinant, mediator, outcome, and indicator of intrinsically motivated behavior is enjoyment. (Deci & Ryan, 1985; Scanlan, Carpenter, Simons, Schmidt, & Keeler, 1993; Wankel, 1993). Enjoyment is an inherent aspect of most leisure activities and is considered one of the core emotions linked to intrinsically motivated human behavior (Izard, 1977). For this reason various theories and health promotion models include the construct of enjoyment, either explicitly or implicitly (Ajzen, 1988; Ajzen & Fishbein, 1974; Bandura, 1986; Deci & Ryan, 1985; Vallerand, 2001; Welk, 1999).

The conceptualization and definition of enjoyment as it relates to exercise psychology has been debated (Kimiecik & Harris, 1996; Wankel, 1997), but generally follows ideas presented by Scanlan & Simons (1992) and Wankel (1993). From their perspective enjoyment is a positive affective state either cognitive or physiological involving feelings of pleasure, fun, and liking.

Others have argued that enjoyment is distinct from positive affect and should be defined in terms of Csikzentmihalyi's (1975) concept of flow (Kimiecik & Harris, 1996). While compelling, this argument fails to address findings that flow can occur without enjoyment, and enjoyment can occur without flow. In addition, fun, pleasure, and positive affect are integral parts of people's definition of enjoyment (Wankel, 1997). For the purpose of this study enjoyment will be defined based on the research and framework of Scanlan and Simons (1992) and Wankel (1993).

Researchers examining motivation to be physically active have related enjoyment to levels of physical activity and adherence to exercise programs. Enjoyment is often cited as an important correlate of physical activity (Dishman & Sallis, 1994; Sallis, 1994; Sallis, Prochaska, & Taylor, 2000) and significant correlations between enjoyment and physical activity behavior have been reported for children and adults using cross-sectional and longitudinal data (DiLorenzo, Stucky-Ropp, Vander Wal, & Gotham, 1998; Garcia & King, 1991; Rovniak, Anderson, Winett, & Stephens, 2002; Salmon, Owen, Crawford, Bauman, & Sallis, 2003; Stucky-Ropp & DiLoenzo, 1993; Thuot, 1995). Evidence examining the effect of enjoyment on change in physical activity as a result of an intervention is, however, limited (Dishman et al., 2005; Lewis, Marcus, Pate, & Dunn, 2002).

It is possible that change in enjoyment of physical activity has little, or no significant, effect on change in physical activity behavior, but few researchers have examined this relationship and fewer have quantified enjoyment adequately. Despite being included as a determinant, or mediator, of behavior change in various theories and health promotion models (social cognitive theory, theory of reasoned action, theory of planned behavior), enjoyment is often operationalized with a single-item or group of items with no previous evidence for reliability or validity.

A number of subscales, stand-alone scales, and single item scales have been used to measure enjoyment of exercise, physical activity, and/or sport. The scales vary in length (1 to 18 items), time frame (enjoyment at this moment to usual level of enjoyment), and type of activity (sport, exercise, general physical activity, or specific activity). These scales are commonly used to look at mediator-moderator relationships (Dishman et al., 2005; Motl et al., 2001), examine intervention outcomes (Calfas et al., 2000; Stevens, Lemmink, van Heuvelen, de Jong, & Rispens, 2003), and compare groups, both experimental and demographic (Moore, Yin, & Gutin, 2004; Ryan, Frederick, Lepes,

Rubio, & Sheldon, 1997; Whitehead & Corbin, 1991). Despite this, very little is known about how the scores and psychometric properties of the different measures of enjoyment compare.

Because there is currently no standardized instrumentation for research related to enjoyment of physical activity, it is difficult to compare results across samples, intervention outcomes, or relationships between enjoyment and physical activity. Limited evidence is available with respect to the factor validity and invariance of most of the scales and less is available about how many of these measures interrelate (Markland & Hardy, 1997; Motl etal., 2001). It may be unlikely that a single measure of enjoyment is accepted (although it would be nice), but the current state of use and development of measures is lacking. Therefore, the purpose of these studies was to compare and extend the construct validity evidence for several measures of physical activity enjoyment. Confirmatory factor analysis was used to examine the factor validity, invariance, and convergence of five enjoyment measures. If the hypothesized factor structure failed to fit the data, exploratory analyses were used to determine a model which was better able to account for the covariance among the items. In addition, each measure of enjoyment was correlated with several estimates of physical activity, perceived competence and extrinsic motives for physical activity.

Statement of problem

Despite indications that enjoyment of physical activity is an important determinant of participation in physical activities, exercise adherence, and a significant moderator of the effect of interventions designed to increase physical activity, limited work has been done to confirm the factor structure, test measurement invariance, or compare scales designed to measure enjoyment of physical activity.

Objectives

The objectives of this research study are to:

1. Test original factor structure of Physical Activity Enjoyment Scale (PACES) in a sample of adult college students and compare to factor structure recently found to fit a modified version of the Physical Activity Enjoyment Scale (PACES-M).
2. Test factorial invariance of the PACES and PACES-M between gender (male and female) and race (Black and White).
3. Compare factor structure and psychometrics of original PACES with PACES-M.
4. Test factor structure and invariance (gender and race) of Groningen Enjoyment Questionnaire (GEQ), enjoyment/interest scale from Intrinsic Motivation Inventory (IMI-joy), and the enjoyment scale from the Motivation for Physical Activity Measure-Revised (MPAM-R-joy).
5. Extend evidence for, or against, the interpretations drawn from enjoyment scores by: (1) Comparing item content, internal consistency reliability, and correlations among enjoyment scales, (2) Comparing the relationships of the various scales to measures of physical activity behavior, perceived physical activity competence, and extrinsic motivations for physical activity, and (3) Testing the convergence of enjoyment scores with a multi-method CFA analysis. This will include comparing three models representing the covariance among all items from the five enjoyment scales.

Delimitations of the study

The research study was delimited to:

1. Five measures of physical activity enjoyment.
2. College students aged 18-24.
3. Comparisons of enjoyment to two measures of perceived competence, three extrinsic motives, and three physical activity questionnaires.
4. Data collected by Web-based survey.

Limitations of the study

        The following limitations are acknowledged as possible threats to the external validity for the study. The measures of enjoyment compared in this study were chosen based on their use in the physical activity literature, current state of development, and item content. All measures of enjoyment could not be included and, although their form and content are very similar to those used; results may not directly apply to measures of enjoyment not included in the study. The results may only generalize to college student aged 18-24. In order to keep the length of the questionnaire manageable a limited number of external criteria were included.

Definition of terms

1. Enjoyment of physical activity - a positive affective state either cognitive or physiological, involving feelings of pleasure, fun, and liking associated with completion or performance of physical activities.
2. Construct – observable trait that can not be directly measured. Often indicated by a set of items thought to measure that trait.
3. Construct validation - the accumulation and integration of evidence that supports, or does not, the use of a test score in a particular situation.
4. Factorial Validity – the extent responses to a set of items conform to the theory underlying item and construct development.
5. Measurement Invariance – the extent to which the relationships among items used to indicate a construct are equivalent across different groups of people (e.g. men and women) or over time.

CHAPTER 2

REVIEW OF RELATED LITERATURE

This chapter is divided into the following sections: (1) The construct of enjoyment in sport, exercise, and physical activity, (2) Enjoyment as a correlate of physical activity behavior and behavior change, (3) Measures of physical activity enjoyment, and (4) Construct validation and confirmatory factor analysis. The purpose of this chapter is to provide background concerning the construct of enjoyment within the sport, physical activity, and exercise literature and to review its current state of measurement. Attention is also paid to general instrument validation with focus on confirmatory factor analysis techniques.

The Construct of Enjoyment in Sport, Exercise, and Physical Activity

Despite the well established benefits of a physically active lifestyle, Americans are generally inactive (USDHHS, 1996). Because of this, countless research studies are conducted and millions of dollars are spent each year in an attempt to understand and change physical activity/exercise behavior ("Cooper Physical Activity Interventions Conference", 1998). In order to understand participation, or lack of participation, in physical activities, exercise, and sport it is important to first be familiar with what motivates people to be physically active.

For centuries, motivation theorists have tried to explain human behavior. From needs and drives to cognitive evaluation and self-efficacy, many have attempted to explain the complexity of what prompts people to act and the direction and persistence of their actions (Deci & Ryan, 1985; Petri & Govern, 2004). Early theories of human motivation took a mechanistic, physiological drive, based approach to explain behavior. These theories focus on behavior as interaction between physiology and environment. Ideas from various psychologists (de Charms, 1968; White, 1959) and

7

the inability of mechanistic theories to fully explain human behavior lead to more organismic theoretical approaches. These theories viewed people as active in their pursuit of needs (both physiological and psychological). From the organismic perspective motivation is a function of both extrinsic and intrinsic sources (Deci and Ryan, 1985; Petri & Govern, 2004).

As defined by Ryan and Deci (2000a) intrinsic motivation is "the inherent tendency to seek out novelty and challenges, to extend and exercise one's capacities, to explore, and to learn". Intrinsic motivation directs and is an important (some would argue *the* most important) determinant of human behavior. Behavior is rarely sustained without intrinsic motivation (Deci, 1975; Deci & Ryan, 1985; Iso-Ahola & St. Clair, 2000; Wankel, 1993; Whitehead, 1993). For example, Mullan and Markland (1997) examined differences in extrinsic and intrinsic motives for exercise across four stages of change for exercise behavior. They found that regular exercisers, those in the maintenance stage, had higher intrinsic motivation compared to individuals in the precontemplation, contemplation, and preparation stages. The difference in intrinsic motivation across stages ranged from 1.5 to 0.50 standard deviations and was similar for men and women. Similarly, when intrinsic motivation for sport participation wanes, dropout is much more likely to occur (Gould & Horn, 1984).

Enjoyment/interest is more related to intrinsic, compared to extrinsic, motivation (Deci, 1975; Deci & Ryan, 1985; White, 1959) and is considered one of the core emotions linked to intrinsically motivated human behavior (Izard, 1977). The conceptualization and definition of enjoyment as it relates to exercise psychology has been debated (Kimiecik & Harris, 1996; Wankel, 1997), but generally follows ideas presented by Scanlan and Simons (1992) and Wankel (1993). From their perspective enjoyment is: "A positive emotion, a positive affect state. It may be homeostatic in nature, resulting from a satiation of biological needs, or growth oriented, involving a cognitive dimension focusing on the perception of successfully applying one's skills to meet environmental

challenges." (Wankel, 1993) and "A positive affective response to the sport experience that reflects feelings and/or perceptions such as pleasure, liking, and experienced fun." (Scanlan & Simons, 1992).

Kimiecik and Harris (1996) argue that enjoyment is distinct from positive affect and should be defined in terms of Csikzentmihalyi's concept of flow (Csikszentmihalyi, 1975, 1990). They do not think enjoyment can be affect, perception, and cognition. They suggest redefining enjoyment as flow citing distinctions between what they consider enjoyment and constructs such as affect, attitudes, pleasure, and intrinsic motivation. Their argument focuses on the difference between constructs often misrepresented as, or equated to, enjoyment (affect, attitude, pleasure, and intrinsic motivation) and redefining enjoyment as flow. For them, enjoyment is process, an end in and of itself. Positive affect is the product, and although related, positive affect, enjoyment, and intrinsic motivation are distinct.

While this argument is compelling, it fails to comment on (as Wankel (1997) points out) "flow" as a construct in and of itself. Csikszentmihalyi (1975; 1990) developed a distinct definition and a sound framework from which to study flow. Flow is defined as an optimal experience occurring when the challenge of an activity is equal or slightly greater than one's ability. It is typically associated with intense concentration, a melding of action and awareness, goal attainment, performance feedback, and distortion of time (Csikszentmihalyi, 1975). Although related, flow is different from enjoyment. In a sense, enjoyment is a characteristic of a flow experience. Ratings of enjoyment have been found to be significantly higher during flow experiences for individuals participating in tennis, basketball, and golf for recreation (Stein, Kimiecik, Daniels, & Jackson, 1995). But flow can occur without enjoyment, and enjoyment can occur without flow (Wankel, 1997). In interviews with adults and children regarding sport and exercise motivation researchers consistently

find fun, pleasure, and positive affect as integral parts of people's definition of enjoyment (Wankel, 1997).

Enjoyment is an inherent aspect of most leisure activities and is considered a determinant, mediator, outcome, and indicator of intrinsically motivated behavior. (Deci and Ryan, 1985; Scanlan et al., 1993; Wankel, 1993). For this reason various theories and health promotion models include the construct of enjoyment, either explicitly or implicitly, when explaining behavior. Although not explicitly defined in Social Cognitive theory (Bandura, 1986), enjoyment could be considered an outcome or a factor that enhances the value of other outcomes. From this perspective, feelings of enjoyment either alone, or in connection, with another outcome, are part of the cognitive process that determines behavior. In the theories of reasoned action and planned behavior (Ajzen, 1988; Ajzen & Fishbein, 1974) enjoyment of a particular activity should enhance ones attitude toward that activity increasing intentions to participate. In these models intentions are the primary mediator of behavior. A person's attitudes reflect feelings about consequences of a behavior. In the case of physical activity, enjoyment of a specific physical activity will improve attitudes toward that activity and ultimately participation (Hagger, Chatzisarantis, & Biddle, 2002). Although change in physical activity enjoyment has not been studied in individuals that move from lower to higher stages of exercise change, results from cross-sectional studies indicate that enjoyment may influence movement from one stage to the next (Matsumoto & Takenaka, 2004; Mullan & Markland, 1997).

Several theories of behavior motivation include the hypothesis that enjoyment of a particular activity is a primary indicator, or outcome, of participation in that behavior. Self-Determination and cognitive evaluation theory posit enjoyment as a primary indicator of intrinsic motivation (Deci, 1975; Deci & Ryan, 1985), while the sports commitment model includes enjoyment as a direct influence on commitment (Scanlan et al., 1993). In addition, enjoyment is included as a primary outcome in Vallerand's (2001) hierarchical model of intrinsic and extrinsic motivation in sport and

10

exercise. This hierarchy includes global, contextual, and situational factors, mediators, motivation, and consequences with reciprocal influences. The model including enjoyment as an outcome has been support in fitness participants followed over a two month period (Vallerand, 2001).

Although intrinsically motivated activities are most associated with enjoyment, researchers examining sources of sport and exercise enjoyment have reported both intrinsic and extrinsic factors underlying enjoyment of a particular activity (Csikszentmihalyi, 1975; Scanlan, Stein, & Ravizza, 1989; Wankel & Kreisel, 1985). Participants in studies examining sources of enjoyment vary in age and activity mode but consistently report comparison to others/competition, personal accomplishment, doing the activity/movement, experience of the activity, use of skills, mental distraction, and friendship/social time as important sources of enjoyment. It is difficult to draw solid conclusions from this research, due to the differences in age, the variety of activity modes, and data collection across studies. Male youth sports participants do, however, rank competition or measuring self against others as a more important source of enjoyment than adults participating in individual physical activities (Csikszentmihalyi, 1975; Heck & Kimiecik, 1993; Scanlan & Simons, 1992; Scanlan et al., 1989; Wankel & Kreisel, 1985).

Enjoyment as a Correlate of Physical Activity Behavior

The relationship between enjoyment and physical activity and its role in changing physical activity behavior have not been adequately studied. While some cite enjoyment as an important correlate of physical activity in adults and children (Dishman & Sallis, 1994; Sallis, 1994; Sallis, Prochaska, Taylor, & Geraci, 2000) others report no relationship (De Bourdeaudhuij & Sallis, 2002; DiLorenzo et al., 1998; Lewis et al., 2002; Neumark-Sztainer, Story, Hannan, Tharp, & Rex, 2003). Conclusions about the association are, however, difficult to compare due to differences in the measures of enjoyment and physical activity used in the various studies.

In order to better understand the relationship between physical activity and enjoyment a literature search was conducted to identify studies that included a measure of enjoyment and physical activity. Eighteen studies were located and sixteen provided information about the correlation between the two variables. Measures of enjoyment ranged from an 18-item scale with some evidence of reliability and validity to single items with no previous support. Measures of activity participation included established measures of physical activity, estimates of sport participation, and fitness measures (e.g. mile run times). A complete list of the studies, sample sizes, measures of enjoyment and activity, and correlations between physical activity and enjoyment can be found in Table 2.1. From the sixteen studies 40 correlation coefficients were obtained. Each correlation coefficient was converted to a Fisher's z before the weighted average was computed. The sample consisted of 9,313 people from 16 studies. The mean correlation, weighted for sample size, between enjoyment and physical activity was 0.22 (95% CI = -0.024 – 0.444), and ranged from -0.12 to 0.42. Although small, if expressed as the binomial effect size display this effect is equivalent to having above average physical activity levels in 61% of people with above average enjoyment scores, but only 38% among people who have below average enjoyment scores.

For youth, enjoyment of physical activity, sport, and physical education have consistently been cited as an underlying reason for participation. (Prochaska, Sallis, Slymen, McKenzie, 2003; Motl et al., 2001; Sallis et al. 1999; Wankel & Kreisel, 1985). Bungum, Dowda, Weston, Trost, Pate (2000) examined correlates of physical activity measured by the previous day physical activity recall in 520 adolescents. Physical activity was quantified as moderate to vigorous (>3 METs) or vigorous (>6 METs). Enjoyment of free time physical activity was estimated with a 3-item bipolar scale. Stability and internal consistency reliability coefficients were above 0.85 for this measure. Mean enjoyment rating did not differ between men and women and similar correlations were reported between enjoyment and vigorous physical activity for males (r = 0.29) and females (r = 0.25). The

correlations between moderate to vigorous physical activity and enjoyment were slightly lower for females (0.14) compared to males (0.25). In another study of psychosocial determinants of physical activity in boys and girls in the 4th through 12th grade, enjoyment of physical education and parental enjoyment of physical activity were used to predict a physical activity index score (Sallis et al., 1999). Enjoyment of physical education was found to be a significant predictor of physical activity for boys across all grade levels (partial r = 0.15), and for girls in grades 7-12 (partial r = 0.25). The partial r's represented the correlation between enjoyment and physical activity after controlling for various demographic, barriers, social, and environmental variables. Parent enjoyment of physical activity was not reported as a significant predictor in any group.

In the competitive sport setting, "fun" and "enjoyment" are often given as reasons for participation (Gill, Gross, & Huddleston, 1983; Scanlan & Lewthwaite, 1986), while dropout is often associated with lack of fun or enjoyment (Gould & Horn, 1984; Salguero, Gonzalez-Boto, Tuero, & Marquez, 2003). In addition, a number of researchers have reported similar reasons for participation in exercise, general physical activity, and recreational sports for adults. In several studies, utilizing focus groups to determine barriers and motivational factors in African American women, investigators reported enjoyment as a primary topic mentioned in relation to being, or becoming, physically active (Nies, Vollman, & Cook, 1999; Wilcox, 2002; Wilcox, Richter, Henderson, Greaney, & Ainsworth, 2002; Young, Gittelsohn, Charleston, Felix-Aaron, & Appel, 2001). Enjoying the exercise itself, or some aspect of the exercise, was identified as a major motivator to maintain activity (Wilcox et al., 2002) as well as a motivation for sedentary women to become more active (Young et al., 2001).

In addition to the qualitative evidence from focus groups, several researchers have attempted to quantified reasons for participation in physical activity and sport. In a study of 422 adults, Ebbeck, Gibbons, and Loken-Dahle (1995) found that enjoyment was the second most important

reason (slightly behind fitness) for participating in a given physical activity. While, Thout (1995) reported that enjoyment was rated the most important reason for continued recreational sports participation after college. Similarly, the most important reason for discontinuing sport participation was "loss of interest".

There is evidence that differences in motivation may exists between individuals participating in sport and those choosing conditioning exercises. For example, Frederick and Ryan (1993) found that interest and enjoyment motivated people who participated in individual sports while those involved in fitness activities were more motivated by physical appearance.

Although limited, there is also evidence that enjoyment can be a significant determinant of physical activity in older adults (Dergance et al., 2003; Johnson & Heller, 1998; McAuley, Jerome, Elavsky, Marquez, & Ramsey, 2003; Stevens, Lemmink, de Greef, & Rispens, 2000). The Groningen active living model was developed as a guide for physical activity promotion among sedentary older adults in The Netherlands (Stevens et. al., 2000). Preliminary cross-sectional data from this project have been used to indicated that enjoyment is significantly higher at progressively higher stages of exercise change (Precontemplation < Contemplation/Preparation < Action/Maintenance). In the 392 sedentary older adults measured during recruitment for the Groningen intervention, enjoyment remained a significant discriminator between exercise stages of change when modeled with self-efficacy, social support, and perceived physical ability. In another study examining participation in a home-based exercise program, benefits and barriers to exercise were measured in 281 adults during hospitalization for myocardial infarction or angina, six weeks after discharge, and six months post discharge. Lack of enjoyment was found to be a significant indicator of individuals who did not adhere to the exercise recommendations they received (Johnson & Heller, 1998).

Few prospective studies have been conducted to investigate the effect of enjoyment on future activity or the influence of enjoyment on physical activity intervention outcomes (DiLorenzo

14

et al., 1998; Dishman et al., 2005; Garcia & King, 1991; Lewis et al., 2002; Trost et al., 1997). Overall, conclusions with respect to the effect of enjoyment on changes in physical activity are mixed. Rovniak et al. (2002) reported significant correlation between enjoyment and physical activity measured 8-weeks later, while other researchers have found that enjoyment during the first 6-months of an exercise program does not predict adherence during the following 6 months (Garcia & King, 1991). In youth, physical activity enjoyment in fifth grade was not found to be a significant predictor of physical activity participation in eight grade when modeled with various social learning and behavioral variables (DiLorenzo et al., 1998), while "liking school PE" in fifth grade has been shown to be a significant predictor of vigorous physical activity in sixth grade (Trost et al., 1997).

Despite the seeming importance, measures of physical activity enjoyment have rarely been included in intervention research (Lewis et al., 2002). Recently, physical activity enjoyment measured with a modified version of the Physical Activity Enjoyment Scale was found to mediate the effects of a school-based intervention (Dishman et al., 2005). In this study, structural equation modeling was used to examine the effects of a year long intervention designed to increase physical activity in adolescent girls. Their results indicated that the small intervention effects on vigorous and moderate to vigorous physical activity were mediated by "factors influencing enjoyment of physical activity" and physical activity enjoyment. Although small, the intervention effect would be meaningful from a population perspective. Also highlighted was the need to target enjoyment during physical activity interventions and to determine factors which will enhance enjoyment of physical activities. If consided in conjunction with the sport and intrinsic motivation literature, emphasis should be placed on competence, control, parental expectations, ability-task matching, and self-esteem (Deci & Ryan, 1985; Scalan & Simons, 1992).

Researchers investigating motivations to be physically active have found positive relationships between enjoyment of physical activity and levels of physical activity, adherence to

15

exercise programs, and intervention outcomes. While, significant small to moderate correlations between enjoyment and physical activity behavior have been reported for children and adults, using cross-sectional and longitudinal data, and in sport and general activity settings the true "effect" of enjoyment on physical activity has not been clearly delineated.   It is possible that change in enjoyment of physical activity has little, or no, effect on change in physical activity behavior, but few studies have examined this relationship and fewer have quantified enjoyment adequately.

Measurement of Physical Activity Enjoyment

The measures of enjoyment compared in this study were chosen based on their use in the physical activity literature, current state of development, and item content. Five measures of enjoyment were selected for evaluation. The scales included the Physical Activity Enjoyment Scale (PACES; Kendzierski & Decarlo, 1991), a modified version of the PACES (PACES-M) developed by Motl et al (2001), the Interest/Enjoyment scale from the Intrinsic Motivation Inventory (IMI-joy; (McAuley, Duncan, & Tammen, 1989; Ryan, 1982), the enjoyment scale from the revised Motivation for Physical Activity Measure (MPAMR-joy; Ryan et al., 1997), and the Groningen Enjoyment Questionnaire (GEQ; Stevens, Moget, de Greef, Lemmink, & Rispens, 2000). These scales, which are commonly used to measure enjoyment of exercise and physical activity, were developed based on studies of sport motivation and performance, intrinsic motivation, and/or flow theory.

The first scale selected, PACES, was a stand alone scale designed specifically to examine enjoyment of physical activities. The PACES was developed from 39 items that were created based on literature reviews, synonyms and antonyms, personal experience, and interviews with individuals about their attitudes toward physical activity (Kendzierski & Decarlo, 1991). Although no specific definition of enjoyment was presented, item content corresponds to several of the current definitions of enjoyment (Kendzierski & Decarlo, 1991; Kimiecik & Harris, 1996; Scanlan et al.,

1993; Wankel, 1993). The 39 items were rated by three exercise adherence experts, 16 were selected and 3 additional items were suggested. The 19-item scale was administered to 30 undergraduates after completing their workout on a stationary exercise bike. One item had a correlation with the total scale score (often called item-total correlation) that was less than 0.30 and was dropped. The remaining 18 items were then administered to 33 students after an abdominal machine workout. Item-total correlations ranged from 0.45 to 0.87. The internal consistency reliability was 0.93.

After scale development, two additional studies were completed by Kendzierski and Decarlo (1991) to provide validity evidence for PACES scores. In the first study, 37 college undergraduates rode an exercise bike for 20 minutes under two experimental conditions. Upon completion of each session they filled out the PACES. One ride was completed in a sterile lab environment with no decorations; this condition was intended to invoke boredom, an un-enjoyable state. The other ride was completed in the same lab environment but participants were allowed to listen to the music of their choice. As expected PACES scores were significantly higher during the external focus (music) condition. In the second study, PACES score was used to predict exercise choice. College undergraduates (n = 34) completed two 15-minute exercise sessions, one on an exercise bike, and one jogging on a mini-trampoline. The PACES was completed after each session. Participants were then brought back for one final session of their choice. Using PACES scores researchers were able to correctly identify the type of exercise chosen by 23 of the 37 (62.2%) participants (Kendzierski & Decarlo, 1991).

The final version of the original PACES consisted of 18-items with a 7-point bipolar rating scale designed to assess enjoyment of a given physical activity. Kendzierski and Decarlo (1991) reported that simply changing the instructions to include the situation or activity being studied is acceptable. Similar values of internal consistency reliability have been reported in adults and children (0.90 – 0.93; Crocker, Bouffard, & Gessaroli, 1995; Kendzierski & Decarlo, 1991). The factor

structure originally hypothesized (single enjoyment factor) has only been tested in children and adolescents (Crocker et al., 1995; Schleppe, 1993). Crocker et al. (1995) tested the factor structure of the original PACES in a sample of sports camp participants (N = 279; mean age = 14.4 years old) competing in track and field, basketball, tennis, and soccer. The researchers reported that mean PACES scores did not differ between males and females, but that the proposed unidimensional factor structure of the PACES did not adequately fit their data (χ^2 (135)= 328.11, GFI = .87, RMSR = .055). In their discussion, they noted several items that may be problematic ("frustrated" and "stimulating") and questioned the unidimensionality of the scale. Although not specifically raised, the inclusion of both males and females and reports of enjoyment from multiple physical activities may make it difficult to clearly assess the lack of model fit reported.

Recently the PACES has been modified for use with adolescent girls involved with a physical activity intervention (Motl et al., 2001). The modified version of the PACES (PACES-M) was altered in several ways. The primary changes from the original PACES were to the rating scale and item wording. The 7-point bipolar rating scale was replaced with single statements rated on a 5-point Likert-type scale. The wording of several items was also changed to enhance understanding. Wording changes were based on focus group sessions with adolescent girls, but were not specifically reported (Motl et al., 2001). In addition, two items were dropped. One due to similarities in content with items already within the scale and the other was found to be irrelevant to enjoyment of adolescent girls.

The items from the PACES and PACES-M are listed in Table 2.2. Of the sixteen items on the PACES-M nine directly match with statements on the PACES. Of the other seven items, six partially match with PACES items and one is not on the PACES. It is interesting to note that comprehension was cited as a reason for the changes made to the PACES-M, but the PACES has been used successfully in a sample of 9 to 17 year old boys and girls with no changes to item

wording and no reported difficulty with comprehension (Crocker et al., 1995).

The factor validity of the PACES-M has been supported in adolescent girls (Motl et al., 2001) and third grade students (Moore et al., 2004). Similar to confirmatory results for the PACES, a single enjoyment factor was not found to adequately represent the relationships among PACES-M items (Motl et al., 2001). Based on exploratory analysis the researchers concluded that a model including a single enjoyment factor and a multidimensional method effect for item wording, positive item uniquenesses allowed to correlate, fit the data best (CFI =0.98, RMSEA = 0.045). This factor structure has been supported in a separate cross validation sample of adolescent girls (Motl et al., 2001) and a sample of young boys and girls (Moore et al., 2004). Although a good fitting model was developed, generalization to the original version of the PACES is limited due to the changes made to the rating scale and items. The factorial invariance of the single factor correlated uniquenesses model for the PACES-M has also been support between black and white adolescent girls (Motl et al., 2001).

Along with support for the factor structure and invariance of the PACES-M, additional evidence for construct validity has been reported (Dishman et al., 2005; Motl et al., 2001). Structural equation models including paths from "factors influencing enjoyment" to PACES-M and from PACES-M to measures of physical activity and sports participation have been tested. In both cross-sectional and longitudinal analyses, significant paths values and model fit supported the construct validity of the PACES-M (Dishman et al., 2005; Motl et al., 2001).

To date, no one has evaluated the factor structure of the original PACES scale in an adult sample, no comparison has been made between the PACES and PACES-M, and no tests of measurement invariance have been conducted for gender or race in adults.

The Groningen Enjoyment Questionnaire (GEQ) was developed as part of an intervention intended to increase leisure-time physical activity in older adults (Stevens et al., 2000). It includes 10

items rated on a 5-point Likert-scale and was designed to be a short, easy to administer, unidimensional measure of enjoyment. The definition of enjoyment which guided instrument development was based on Csikzentmihalyi's (1975) concept of autotelic experience (one of the nine dimensions of flow). Based on this definition and a search of the literature 30 items were formed during scale development. Exploratory factor analysis was used to identify the ten items selected for the final scale. Factor loadings from the exploratory analysis ranged from 0.58 to 0.86. Coefficient alpha (0.88) and test-retest reliability of the total score (0.84) indicated acceptable internal consistency and stability over a two-week period.

Evidence for the validity of scores from the GEQ was presented in the initial findings from the Groningen active living model. This model was developed to explain and guide a behavioral change intervention for older adults. The intervention is designed to increase perceived fitness, social support, and self-efficacy which should increase enjoyment of physical activity leading to increased program adherence and leisure time physical activity. Cross-sectional data from 329 sedentary adults (55-65 years old) supported using GEQ scores to discriminate across stage of change for exercise (Stevens et al., 2000). In addition, GEQ scores were found to significantly correlate with perceived ability (0.31 – 0.44), self-efficacy (0.45 – 0.55), and social support (0.17 – 0.29). Of this recruited sample, 96 then completed an 18-month Groningen active living model intervention. Enjoyment scores were found to significantly discriminate between adherers and drop-outs at 6- and 18-months of the intervention (Stevens et al., 2003).

Development of the Intrinsic Motivation Inventory (IMI) is credited to Ryan (1982), but the instrument has been more thoroughly investigated with respect to exercise and sport by McAuley and colleagues (McAuley, Duncan, & Tammen, 1989; McAuley, Wraith, & Duncan, 1991). The instrument consists of 27-items designed to measure five underlying dimensions of intrinsic motivation (interest-enjoyment, perceived competence, effort, pressure-tension, and perceived

choice). The interest/enjoyment scale contains 7-items rated on a 7-point scale ranging from "not at all true" (1), "somewhat true" (4), and "very true" (7). McAuley et al. (1991) reported that similar items within a subscale, or unneeded subscales, can be removed without changing the internal consistency reliability or factor structure of the remaining items. Item wording can also be changed to reflect intrinsic motivation in a variety of situations.

Different numbers of items (5 or 7) have been used to represent the enjoyment/interest subscale of the IMI. Each version has been found to have acceptable internal consistency reliability (0.78 – 0.92; Markland & Hardy, 1997; McAuley et al. 1989;1991; Tsigilis & Theodosious 2003) and stability reliability (0.86; Tsigilis & Theodosiou, 2003). Conflicting findings have raised questions regarding the factor structure of the IMI in sport and exercise settings (Markland & Hardy, 1997; McAuley et al., 1989; McAuley et al., 1991). Using a version of the IMI reflecting motivation toward a basketball game Mcauley et al. (1989) examined the properties and factor structure of the scale in a competitive sport setting. During the study 116 undergraduate students competed in a modified shooting task, based on the classic basketball game "HORSE". Upon completion an 18-item, four subscale version of the IMI was completed. Confirmatory factor analysis was used to test several competing structural models of the IMI. Although factor loadings were significant and factor correlations were as expected, the overall fit of the model to the data did not support the second- or first-order factor structure of the scale (CFI = 0.81, RMSR = 0.112).

In contrast, a first-order correlated factors model has been found to fit data collected from 265 participants in an aerobic exercise class. In this study, a 21-item, five subscale version of the IMI worded to measure motivation in aerobic exercise was administered to beginner, intermediate, and advanced participants in a 10-week aerobic dance class (McAuley et al., 1991). The fit of a five correlated factors model was good (CFI = 0.94, RMSR = 0.057), while a second-order model was concluded to be acceptable (CFI = 0.92, RMSR = 0.075).

The interest/enjoyment subscale (IMI-joy) is considered the self-report measure of intrinsic motivation in the IMI. Acceptable internal consistency reliabilities have been reported for 3, 5, and 7 item versions of the IMI-joy subscale (range 0.78 – 0.92; Markland & Hardy, 1997; McAuley et al. 1989;1991; Tsigilis & Theodosious 2003). In two studies by McAuley and colleagues (1989; 1991) and a study by Markland and Hardy (1997) the item reflecting "thinking about enjoyment during the activity" was found to be problematic (small factor loading and R^2). Results were often reported with this item excluded.

The final enjoyment scale reviewed is from the revised Motivations for Physical Activity Measure (MPAMR; Ryan et al., 1997). The original version of the MPAM contains 23 items tapping three motives for physical activity; body-related, competence, and interest/enjoyment (Frederick & Ryan, 1993). The factor structure was supported in undergraduate college students (n = 150) and an adult population (n = 376; mean age 39). The interest/enjoyment subscale from the original MPAM and MPAMR are identical except that one item ("Because I like the excitement of participation") was added to the revised version.

The MPAMR was developed to differentiate between appearance and fitness goals contained within the body-related motives scale in the original MPAM. In addition, a social motives scale was added. The separation of appearance and fitness goals and inclusion of social motive is though to provide more comprehensive coverage of intrinsic and extrinsic factors which are related to participation (Ryan et al., 1997). The scale was revised based on the findings of two pilot studies. During the revision, the scale was expanded to 30 items and 5 factors (fitness, appearance, competence, enjoyment, and social). Each item is rated on a 7-point Likert-type scale ranging from "not at all true for me" (1) to "very true for me" (7).

Estimates of internal consistency reliability for the MPAMR interest/enjoyment subscale have consistently been around 0.90. Interest/enjoyment scores have also been shown to correlate

with energy expenditure (r = .30), length of workout (r = 0.23), attendance (r = 0.52), and dropout (r = -0.43). Mean interest/enjoyment scores have not been found to differ between men and women, but significantly higher scores have been noted for adherers to an exercise program compared to non-adherers and for individuals participating in sports compared to fitness exercisers (Ryan et al., 1997).

The factor structure of the MPAM-R has been supported by exploratory factor analysis. Using data from 155 new members at a university fitness center Ryan et al. (1997) found a five factor solution similar to the one proposed. The enjoyment factor accounted for more variance than other subscales with factor loadings ranging from .49 to .85. Two of the seven enjoyment items did, however, cross load on the competence factor.

The scales reviewed have been used as indicators of enjoyment of physical activity and exercise in a variety of situations. Because there is currently no standardized instrumentation for research related to enjoyment of physical activity or exercise behavior it is difficult to compare results across samples, intervention outcomes, or relationships between enjoyment and physical activity. There is limited evidence for the reliability and construct validity of most of the scales and less support for the factorial validity and invariance. No studies were found that specifically examined the convergent and divergent validity of enjoyment measures. One study was located that included both the IMI and PACES (Ryska, 2003). The correlation between PACES and IMI interest/enjoyment subscale was 0.461. This scale was also reported to be the most important IMI subscale in a prediction of PACES score.

Instruments used to measure enjoyment of physical activity differ in length, context, and to some extent content, and while it may not be reasonable to expect the use of a single accepted measure of enjoyment the current state of use and development are lacking. This along with suggestions calling for comparisons of similarly named constructs used in different behavioral

theories (Nigg, Allegrante, & Ory, 2002) seems to indicate that more detailed investigation of physical activity enjoyment measures are needed.

Construct Validity and Confirmatory Factor Analysis

Validity concerns the appropriateness, or accuracy, of inferences made from scores on a test. Test validation is in no way a finite process, rather it involves the accumulation of experimental, statistical, observational, and theoretical evidence supporting the various uses and interpretations of a given test score (Messick, 1989). In psychology, test scores are most often used to represent traits, or constructs, that can not be directly measured. An individual's "ability" on a given construct is inferred from his/her responses to items representing that trait. Although validity is, and should be, considered a unitary concept, various sources of evidence are often cited as important when evaluating test scores. Three of the many sources of evidence, often falling under the omnibus heading - construct validity, are content, structural, and external relationships (AERA, 1999; Messick, 1995). Each gives credibility to the use of a test, increasing the confidence researchers can place on interpretation from a particular score.

Content related evidence is the extent to which the items on a test are relevant and representative of the domain, or construct of interest (Wood, 1989). Conceptually, there is some controversy over the inclusion of content related evidence as an area of test validation (Messick, 1975; Wood, 1989). In essence, the debate concerns the focus on items, instead of score interpretations. Although important, some argue that evidence provided by examining content does not reflect an interpretation of scores from a test, rather the relevance and representativeness of individual items. Despite this, the Standards for Educational and Psychological Testing (AERA, 1999) include test content as a primary source of validity evidence for a test.

The accumulation of content related evidence usually occurs during test construction (Wood, 1989). The initial step involves defining the construct, or domain of interest. This definition is then used to create items that are representative of the domain to be measured. Items can then be evaluated by a group of experts. If judged by these experts to be relevant and representative of the construct, the items can begin, or continue, the test validation process.

Structural evidence can be examined in terms of how the relationships among items used to indicate a construct conform to the conceptual framework of the test (AERA, 1999; Messick, 1995). If a test is intended to measure a unitary concept (e.g. enjoyment of physical activity) then items on that test should be highly related because they are all measuring enjoyment. If they are not highly related, or some are more related than others, the unidimensionality of the scale may be questioned. For example, in a test intended to measure enjoyment of physical activity, the variation in responses to each item, should reflect variance in the construct of interest (enjoyment of physical activity) plus some small unrelated random error. Once "enjoyment" is accounted for, any residual variance for a particular item should be unrelated to the residual variance for other items on the test. If, after accounting for "enjoyment", the residual variances for some items are still related it could be concluded that the test does not only measure enjoyment. This is because the structural evidence does not support the conceptual frame work for the test (a single construct - enjoyment of physical activity).

Internal consistency reliabilities are often reported as evidence for the internal structure of a test. A better method for providing this evidence is factor analysis. In essence, factor analysis is a data reduction technique (Bollen, 1989). It helps to identify a small number of latent dimensions that explain the relationships among a larger number of observed variables. In terms of validity evidence, the technique can be used to support, or confirm, hypotheses about the internal structure of a measure. Exploratory factor analysis is often used to select or eliminate items during test

development. In exploratory factor analysis the question is: Can the covariance among these items be explained by one or more latent dimensions? The hope is that the factors that emerge from the data match with the conceptualization of the construct.

In confirmatory factor analysis the inquiry is reversed. The question becomes: Does this hypothesized model adequately explain the covariance among these items? Therefore, confirmatory factor analysis is used to test a priori hypotheses about the relationships among observed variables (Bollen, 1989; Lance & Vandenberg, 2001). This involves fixing and/or freeing sets of parameters that will be used to reproduce the sample covariance matrix. In confirmatory factor analysis, the hypothesis about the underlying structure of a test is judged by comparing observed relationship among items to those estimated from the proposed model of the test.

A number of indices are currently used to asses model fit in confirmatory factor analyses (Hu & Bentler, 1998; Hu & Bentler, 1999). Traditionally, a direct test comparing the sample covariance matrix to the reproduced matrix is evaluated. This chi-square statistic tests badness of fit. A drawback of the chi-square is its sensitivity to sample size. Because large sample sizes are required to obtain stable parameter estimates in confirmatory factor analysis, the chi-square statistic calls for rejection of the hypothesized model when discrepancies between the sample and reproduced covariance matrix are quite small (Bollen, 1989; Joreskog, 1993). For this reason, it is suggested that several other measures of fit be reported (Byrne, 1998; Hu & Bentler, 1998; 1999).

Fit indices are often classified as absolute or incremental. Absolute indices are used to directly estimate how well the model reproduces the covariance among items. The root mean square error of approximation (RMSEA) is a standardized estimation that represents closeness of fit of population data to the model (Steiger & Lind, 1980). It represents the difference between the reproduced covariance matrices and the population covariance matrix per degree of freedom. The RMSEA is widely used and is considered one of the most informative fit criteria (Byrne, 1998).

Values below 0.05 indicate good fit, those around 0.06 reflect close fit, while values up to and around 0.10 denote reasonable fit of the model to the data (Browne & Cudeck, 1993; Hu & Bentler, 1998). The 90% confidence interval (CI) around the RMSEA point estimate should contain 0.06 or zero to indicate the possibility of close or exact fit.

The standardized root mean square residual (SRMSR) is another absolute index of fit. The SRMSR is the standardized average of the discrepancy between the sample covariance matrix and the covariance matrix reproduced from the specified model. In other words, it is an average residual that has been standardized to a correlation metric (Hu & Bentler, 1998; Joreskog, 1993).

Incremental indices are used to judge model fit in comparison to some baseline model (Byrne, 1998; Hu & Bentler, 1998). In most cases the baseline, or null, model includes all observed variables but specifies that they are unrelated. These indices are conceptualized as the proportionate improvement in fit of the target model over the baseline model (Bentler & Bonett, 1980; Hu & Bentler, 1998). Incremental indices can be further classified based on the type of information used in calculating improvement over the null model.

The non-normed fit index (NNFI) and the comparative fit index (CFI) are commonly used incremental fit indices. The CFI and NNFI test the proportionate improvement in fit by comparing the target model to the baseline model using expectations of overall model fit (Bentler & Bonett, 1980; Hu & Bentler, 1999). The NNFI, but not the CFI, is affected by model parsimony (more complex models are penalized). Values for the CFI and NNFI greater than 0.95 have been shown to indicate good fit (Hu & Bentler, 1999), while values greater than 0.90 are considered acceptable, especially with large sample sizes (Akaike, 1987; Bandalos, 1993; Bentler & Bonett, 1980; Browne & Cudeck, 1989; Hu & Bentler, 1998; 1999; Marsh, Hau, & Wen, 2004). Although criteria have been proposed for judging model fit (Hu & Bentler, 1998; 1999) some researchers suggest that strict "cut-off" values for specific indices may not always be appropriate (Byrne, 1998; Hu & Bentler, 1999;

Marsh et al., 2004). For this reason, overall assessment of model fit should be based on model interpretability and the cumulative evidence from all measures of fit.

Another important aspect of validity evidence involves estimating the relationship of the test scores with other variables of interest. The actual validity evidence evolves from how these relationships conform to theoretical hypotheses about the construct (Cronbach & Meehl, 1955; Messick, 1989, 1995). One aspect of this involves developing support for convergence and divergence of constructs. Using enjoyment of physical activity as an example, evidence for convergence could be examined by estimating the relationships among several measures of enjoyment of physical activity. Evidence for divergence would involve estimating the relationships between enjoyment of physical activity and another construct. Campbell and Fiske (1959) proposed the multitrait-multimethod (MTMM) matrix as a method for examining these relationships. By measuring distinct traits, or constructs, with different methods and calculating correlations among those variables convergence, divergence, and method effects can be evaluated. One difficulty with the MTMM analysis is that conclusions drawn from the correlation matrix are subjective. More recently, confirmatory factor analysis techniques have been applied to MTMM data, allowing researchers to tests the significance of convergence, divergence, and method effects within the MTMM matrix (Kenny & Kashy, 1992; Lance, Noble, & Scullen, 2002; Widaman, 1985).

In addition to simple comparisons between tests measuring the same construct and those measuring different constructs, evidence for the validity of score interpretation comes from larger theoretical models. These models involve building, what Cronbach and Meel (1955) referred to as, a nomological network "around" the construct of interest. Evidence for the relationships of the construct within a single or across multiple theoretical models strengthens the interpretations that can be made with respect to a particular test score (AERA, 1999; Messick, 1989, 1995).

Validity is often considered the most important aspect of a test score. Cronbach (1988) and Messick (1995) write of validation as an "evaluation argument". In this sense, researchers are continuously trying to convince themselves that a test score interpretation is accurate. In the end, validation is the accumulation and integration of evidence that supports, or does not, the use of a tests score in a particular situation. Each piece of validity evidence is important and none can stand alone.

Table 2.1 Relationship between measures of physical activity, sport, and fitness and measures of enjoyment

Study	N	Gender	Enjoyment measure	PA/Fit/Sport	r
Bungum et al. (2000)	255	Male	3-I, Enjoy exercise	PDPAR Moderate-vigorous PA	0.250
	265	Female	3-I, Enjoy exercise	PDPAR Moderate-vigorous PA	0.140
	255	Male	3-I, Enjoy exercise	PDPAR Vigorous PA	0.290
	265	Female	3-I, Enjoy exercise	PDPAR Vigorous PA	0.250
Calfas et al. (2000)	101	Female	5-I from PACES	Walking activity	0.180
	121	Females	5-I from PACES	Walking activity	0.170
Davison & Birch (2001)	197	Female	1-I, parent enjoyment	Days per week of activity	0.310
	197	Males	1-I, parent enjoyment	Days per week of activity	0.380
Dishman et al. (2005)	2087	Female	PACES-M	3DPAR Vigorous PA	0.200
	2088	Female	PACES-M	3DPAR Vigorous PA	0.103
Frederick et al. (1993)	232	Both	MPAM	Energy expenditure	0.250
	103	Both	MPAM	Energy expenditure	0.300
	232	Both	MPAM	Hours per week of exercise	0.320
	103	Both	MPAM	Hours per week of exercise	0.320
Frederick et al. (1996)	38	Male	MPAM-R	Day per week of Exercise	0.160
	38	Male	MPAM-R	Hours per week of exercise	0.230
	80	Female	MPAM-R	Day per week of Exercise	0.040
	80	Female	MPAM-R	Hours per week of exercise	0.250
Garcia & King (1991)	74	Both	1-I, 1-10 rating scale	Adherence to exercise progam	0.024
Klentrou et al. (2003)	256	Both	Enjoy PE	VO2max	0.125
Motl et al. (2001)	1797	Female	PACES-M	3DPAR moderate-vigorous PA	0.230
	1797	Female	PACES-M	Sport participation	0.370
Neumark-Sztainer et al. (2003)	201	Female	4-I, Enjoy, fun, interest	Moderate - Vigorous PA	0.070
Prochaska et al. (2003)	414	Both	Enjoy PE	Mile run	0.270
	414	Both	Enjoy PE	Sport participation	0.140
Rovniak et al. (2002)	277	Both	PACES	Stage of exercise change	0.420
	278	Both	PACES	Kcal/Kg/wk	0.350
	279	Both	PACES	# physical activities	0.280
Saelens & Epstein (1999)	39	Both	PACES	PA recall interview	0.250
Sallis et al. (1999)	247	Boys	Enjoy PE	PA Index	0.148
	250	Boys	Enjoy PE	PA Index	0.149
	200	Boys	Enjoy PE	PA Index	0.153
	229	Girls	Enjoy PE	PA Index	0.000
	208	Girls	Enjoy PE	PA Index	0.224
	210	Girls	Enjoy PE	PA Index	0.275
Trost et al. (1997)	92	Boys	Enjoy PE	3DPAR Moderate-Vigorous PA	-0.120
	110	Girls	Enjoy PE	3DPAR Moderate-Vigorous PA	0.210
	92	Boys	Enjoy PE	3DPAR Vigorous PA	0.020
	110	Girls	Enjoy PE	3DPAR Vigorous PA	0.260
Vilhjalmsson et al. 1998	1131	Both	Enjoy PE	Day/hours summary score	0.183

Note: PDPAR = Previous Day Physical Activity Recall; PACES = Physical Activity Enjoyment Scale; MPAM = Motivation for Physical Activity Measure; r = correlation coefficient; PA = Physical Activity

Table 2.2 Comparison of items from Original and Modified Physical Activity Enjoyment Scales

I#	PACES			I#	PACES-M
1	I enjoy it	…	I hate it	1	I enjoy it.
2	I feel bored	…	I feel interested	2	I feel bored.
3	I dislike it	…	I like it	3	I dislike it.
4	I find it pleasurable	…	I find it unpleasurable	?	NO MATCH
5	I am very absorbed in this activity	…	I am not at all absorbed in this activity	*	DROPPED
6	It's no fun at all	…	It's a lot of fun	5	It's no fun at all.
7	I find it energizing	…	I find it tiring	6	It gives me energy.
8	It makes me depressed	…	It makes me happy	7	It makes me depressed.
9	It's very pleasant	…	It's very unpleasant	8	It's very pleasant.
10	I feel good physically while doing it	…	I feel bad physically while doing it	9	My body feels good.
11	It's very invigorating	…	it's not at all invigorating	*	DROPPED
12	I am very frustrated by it	…	I am not at all frustrated by it	12	It frustrates me.
13	It's very gratifying	…	It's not at all gratifying	?	NO MATCH
14	It's very exhilarating	…	It's not at all exhilarating	11	It's very exciting.*
15	It's not at all stimulating	…	It's very stimulating	?	NO MATCH
16	It gives me a strong sense of accomplishment	…	It does not give me any sense of accomplishment at all	14	It gives me a strong feeling of success.
17	It's very refreshing	…	It's not at all refreshing	?	NO MATCH
18	I felt as though I would rather be doing something else	…	I felt as though there was nothing else I would rather be doing	16	I feel as though I would rather be doing something else.
			See # 2 PACES	13	It's not at all interesting.
			See # 6 PACES	10	I find it fun.
				4	I get something out of it.
				15	It feels good.

* judged to be match even though descriptors not exactly the same

CHAPTER 3

FACTOR VALIDITY AND INVARIANCE OF THE ORIGINAL AND MODIFIED

PHYSICAL ACTIVITY ENJOYMENT SCALES[1]

---

[1] Hales, D.P. & Baumgartner, T.A. To be submitted to Journal of Physical Activity and Health

Abstract

Enjoyment is cited as an important correlate of physical activity, but is often measured with a single-item or group of items with no previous evidence for reliability or validity. Two instruments, the Physical Activity Enjoyment Scale (PACES) and the modified PACES (PACES-M) are promising possibilities for measuring enjoyment of physical activity. The purpose of this study was to evaluate the factor structure of the PACES, cross-validate the factor structure of the PACES-M, compare the PACES and PACES-M, and test measurement invariance across gender and race for each scale in a sample of young adults. Participants (N = 1023; mean age 19.60 ± 1.55 years) were recruited from basic physical education classes and the psychology research pool at two universities. Confirmatory factor analysis was used to examine factor validity and invariance for both scales. The results supported the hypothesized factor structure of the PACES-M (CFI = .977; RMSEA = 0.055), but not the PACES (CFI = .844; RMSEA = 0.112). Based on exploratory analyses, a two factor model was found to represent the PACES (CFI = 0.940, RMSEA = 0.078). In addition, the factorial invariance between black and white and male and female participants was supported for both scales. Mean PACES and PACES-M scores were significantly higher for males compared to females. Mean PACES score was found to be significantly higher for white compared to black participants. This was not true for the PACES-M. Factor correlations among the PACES-M and PACES were moderate to high, supporting similarity of measurement but not complete convergence.

Key Words: Confirmatory factor analysis; Enjoyment; Invariance; Physical activity

Introduction

Despite the well established benefits of a physically active lifestyle, Americans are generally

inactive (US Department of Health and Human Services [USDHHS], 1996). Currently between 30%

and 50% of adults get the recommended amount of physical activity, while 25% report no leisure

time physical activity (USDHHS, 2003a; USDHHS, 2004). Alarmingly, low levels of physical activity

are increasing among youth with 40% of girls and 27% of boys not getting the recommended

amount of moderate or vigorous physical activity (USDHHS, 2003b). These low levels of physical

activity, in conjunction with the consistent evidence that physically active individuals are healthier

than their sedentary counterparts (USDHHS, 1996) has lead to the establishment of national

physical activity goals and a multitude of interventions designed to increase physical activity levels.

Physical activity, or health related, interventions are often designed, not only to increase

activity levels, but to target specific psychosocial correlates thought to determine physical activity

behavior. Enjoyment is one factor often cited as an important correlate of physical activity

(DiLorenzo, Stucky-Ropp, Vander Wal, & Gotham, 1998; Dishman & Sallis, 1994; Garcia & King,

1991; Rovniak, Anderson, Winett, & Stephens, 2002; Sallis, 1994; Sallis, Prochaska, & Taylor, 2000;

Salmon, Owen, Crawford, Bauman, & Sallis, 2003; Stucky-Ropp & DiLoenzo, 1993; Thuot, 1995),

influence on adherence to exercise programs (Johnson & Heller, 1998; Ryan, Frederick, Lepes,

Rubio, & Sheldon, 1997), and mediator of intervention effects on changes in physical activity

(Dishman et al., 2005; Lewis, Marcus, Pate, & Dunn, 2002).

Enjoyment has been defined as a positive affective state either cognitive or physiological,

involving feelings of pleasure, fun, and liking (Scanlan & Simons, 1992; Wankel, 1993). Because

enjoyment is an inherent aspect of most freely chosen leisure activities (Scanlan & Simons, 1992;

Wankel, 1993) and is considered one of the core emotions linked to intrinsically motivated human

34

behavior (Izard, 1977) various theories and health promotion models include the construct of enjoyment, either explicitly or implicitly (Ajzen, 1988; Ajzen & Fishbein, 1974; Bandura, 1986; Deci, 1975; Deci & Ryan, 1985; Vallerand, 2001; Welk, 1999). Despite this, enjoyment is often measured with a single-item or group of items with no previous evidence for reliability or validity. Two instruments, the Physical Activity Enjoyment Scale (PACES; Kendzierski & Decarlo, 1991) and the modified PACES (PACES-M; Motl et al., 2001) are promising measures of physical activity enjoyment.

The PACES and PACES-M have been used to measure enjoyment of exercise, sports, and general physical activity in adults and children. Although no definition of enjoyment was specifically stated during development, item content on both scales corresponds to several of the current descriptions of enjoyment in the exercise, sport, and physical activity literature (Kendzierski & Decarlo, 1991; Kimiecik & Harris, 1996; Scanlan, Carpenter, Schmidt, Simons, & Keeler, 1993; Wankel, 1993). In addition, instructions and items content were written so that the scale could be used in a variety of research situations (Kendzierski & Decarlo, 1991; Kimiecik & Harris, 1996; Scanlan et al., 1993; Wankel, 1993). Instructions for the scale have been modified to reflect enjoyment of fitness classes, specific exercises and sports, walking, general physical activity, and sedentary behaviors (Crocker, Bouffard, & Gessaroli, 1995; Motl et al., 2001; Rovniak et al., 2002; Saelens & Epstein, 1999; Wininger, 2002).

The PACES was developed from 39 items that were created from literature reviews, synonyms and antonyms of enjoyment, personal experience, and interviews with individuals about their attitudes toward physical activity (Kendzierski & Decarlo, 1991). Each item was rated by three exercise adherence experts and 16 items were selected for inclusion in the scale. Three additional items were suggested and agreed upon by the experts. These 19 items were then evaluated using two small (n = 34 and 37) samples of college students. The final version of the original PACES consisted

of 18 bipolar items designed to assess enjoyment of a given physical activity. The correlations of the items with the scale total score (item-total correlation) ranged from 0.45 to 0.87. Similar estimates of internal consistency reliability have been reported for adults (0.93; Kendzierski & Decarlo, 1991) and children (0.90; Crocker et al., 1995).

Although developed using college students, the factor structure originally hypothesized (single enjoyment factor) has only been tested in children and adolescents (Crocker et al., 1995; Schleppe, 1993) Exploratory factor analysis results have supported a one-factor solution, confirmatory factor analysis has not. Crocker et al. (1995) tested the factor structure of the original PACES in a sample of sports camp participants (N = 279; mean age 14.4 years) competing in track and field, basketball, tennis, and soccer. They reported that mean PACES scores were similar for males and females, but that the proposed single factor model did not adequately fit their data (χ^2 (135)= 328.11, GFI = .87, RMSR = .055). Despite lack of support for factor validity, scores from the original PACES have accurately predicted activity choice, been shown to mediate mood changes, differentiated between exercise session (music vs. no music), and been used to indicate outcome expectancies. In their discussion, Crocker et al. (1995) noted several items that may be problematic ("frustrated" and "stimulating") and questioned the unidimensionality of the scale. Although not specifically raised by Crocker et al. (1995), the inclusion of both males and females and reports of enjoyment from multiple activities may make it difficult to clearly assess the lack of model fit.

The PACES-M was developed for use with adolescent girls involved in a physical activity intervention. The primary changes from the original PACES were to the rating scale and item wording. The 7-point bipolar scale (PACES) was replaced with single statements rated on a five point Likert-type scale. The wording was also changed for several items to enhance understanding. Wording changes were based on focus group sessions with adolescent girls, but were not specifically reported (Motl et al., 2001). In addition, two items were dropped. One due to similarities in content

with items already within the scale and the other was found to be irrelevant to enjoyment of adolescent girls.

The factor validity of the PACES-M has been supported in adolescent girls (Motl et al., 2001) and third grade students (Moore, Yin, & Gutin, 2004). Similar to confirmatory results for the PACES, a single enjoyment factor was not found to adequately represent the relationships among PACES-M items (Motl et al., 2001). Based on exploratory analyses the researchers concluded that a model including a single enjoyment factor and a multidimensional method effect for item wording, allowing positive item uniquenesses to correlate, fit the data best (CFI =0.98, RMSEA = 0.045). This factor structure has been supported in a separate cross validation sample of adolescent girls (Motl et al., 2001) and a sample of third grade boys and girls (Moore et al., 2004). Although a good fitting model was developed, generalization to the original version of the PACES is limited due to the changes made to the rating scale and items. The factorial invariance of the PACES-M has also been supported between black and white adolescent girls (Motl et al., 2001).

Along with support for the factor structure and invariance of the PACES-M, additional evidence for construct validity has been reported (Dishman et al., 2005; Motl et al., 2001). Structural equation models including paths from "factors influencing enjoyment" to PACES-M and from PACES-M to measures of physical activity and sports participation have been tested. In both cross-sectional and longitudinal analyses, significant paths values and model fit supported the construct validity of the PACES-M (Dishman et al., 2005; Motl et al., 2001).

If PACES and PACES-M scores are used to compared groups or as an influence on, or outcome, of physical activity, it is necessary to first establish the factor validity and invariance of responses to these scales in multiple samples. Factor validity is the extent a set of items, representing a construct, conforms to the theoretical definition of the construct (Bollen, 1989; Messick, 1995). While, factor invariance is the extent to which the relationship among items

generalizes across points of time or between groups of people. Tests of factor invariance are used to examine the comparability of the form and values of parameters within a measurement model (Vandenberg & Lance, 2000). To our knowledge, no researchers have resolved the factor structure of the PACES or tested multi-group invariance. The factor structure and invariance of the PACES-M have been supported, but only in children and adolescent girls. In addition, despite the connection between the scales (one is a modification of the other), and references to one as support for the other, no direct comparisons have been made. Therefore, the purpose of this study was to evaluate the factor structure of the original PACES scale, cross-validate the factor structure of the PACES-M, compare the original and modified versions of the PACES, and test measurement invariance across gender and race in a sample of young adults.

<center>Method</center>

Participants

Participants ($N = 1023$) were recruited from basic physical education classes and the psychology research pool at two large southern universities ($n = 733$; $n = 290$). Students were informed of the study through direct class contact or a listing of available experiments. Psychology research participants received class credit for involvement in the study. The sample had a mean age of 19.60 ± 1.55 years with a racial distribution of 73.9% white, 18.8% black, and 7.3% other. Most participants were freshman ($n=405$), but sophomores ($n = 238$), juniors ($n =204$) and seniors ($n=174$) also completed the survey. Sixty-eight percent of the participants were female and 27.6% were recruited from physical education classes.

Measures

The PACES contains 18 items designed to assess enjoyment of a given physical activity (Kendzierski & Decarlo, 1991). Each item consists of two bipolar statements with a seven point scale defining the continuum between them (e.g. "I dislike it. 1 2 3 4 5 6 7 I like it."). Respondents are asked to rate how they felt at the moment about the physical activities in which they usually participate. A total score is calculated as the sum, or mean, of the item scores. The PACES-M contains 16-items (Motl et al., 2001). Each item consists of a single statement that is rated on a 5-point scale ranging from (1) "Disagree a lot" to (5) "Agree a lot". Changes from the original version are presented in the introduction.

Procedures

A web-based survey was used for data collection. In a pilot study comparing web-based and paper-and-pencil response no significant differences in item or total score for the PACES or PACES-M were found between the two methods of administration. The internal consistency of both scales was greater than 0.90 and similar across paper and computer versions of the survey. The single measure Intraclass R for the PACES (0.73) and PACES-M (0.76) indicate acceptable stability of the scales across the two methods of administration. The web-survey was constructed with help from the campus Survey Research Center. Each survey contained an informed consent section, demographic questions (age, sex, etc.), measures of enjoyment, and a social desirability scale. Each questionnaire (PACES and PACES-M) was displayed as a separate section within the web-survey and replicated to match the layout and response format of the original paper-and-pencil version. The survey took less than 20 minutes to complete.

Once enrolled in the study, students were contacted via email. Each email message contained a brief overview of the study, a hyperlink to the survey web-site, and a personalized passcode.

Passcodes were used to limit access to recruited participants. To enhance response rate, email messages were sent one and two weeks after the initial contact to remind people to complete the survey. All procedures were approved by the institutional review boards at the universities where data collection took place.

Data analysis

Means, standard deviations, measures of distribution, Pearson product moment correlations, and internal consistency reliability coefficients were calculated using SPSS 13.0. Independent samples t-test was used to compare total scores between groups. PRELIS (Joreskog & Sorbom, 1996) was utilized to estimate tests of multivariate normality. Initial tests of the hypothesized factor structure for the PACES and PACES-M were conducted using confirmatory factor analysis.

The hypothesized model for the PACES included a single enjoyment factor indicated by 18 items. The PACES-M model included a single enjoyment factor indicated by 16 items with a method effect represented by correlated uniquenesses among the positively worded items. The uniqueness, or measurement error, for each item is the variance not accounted for by the common factor. When the initial model for the PACES failed to fit the data, the total sample was split into a test sample (n = 409) and a cross-validation sample (n = 614). An exploratory analysis was then conducted using the test sample. The results from the exploratory factor analysis were used to re-specify the model for the PACES. This new model was estimated in the test sample using confirmatory factor analysis. Additional areas of model misfit were identified by examining modification indices, parameter estimates, squared multiple correlations, and residuals. Item content and previous results were also used to guide model modification. The final model from the exploratory analysis was then fit to the data from the cross-validation sample.

To confirm the method effect previously reported for the PACES-M, four models were evaluated and compared. The first model included a single enjoyment factor indicated by 16 items. The next three models have been suggested to assess whether factors representing positive and negative item wording are meaningful or methodological artifact (Marsh, 1996). The second model included two factors, one indicated by the positively worded items and the other indicated by the negative items. The third and fourth models were very similar. Both included a single enjoyment factor indicated by the 16 items and a method effect represented by correlated uniquenesses. In the third model, uniquenesses among the negatively worded items were allowed to correlate. In the fourth model, correlated uniquenesses for the positively worded items were estimated.

Exploratory factor analysis. The exploratory factor analysis was conducted using principal axis factoring with oblique rotation. This common, or principal, factor analysis technique used squared multiple correlations to produce iterated estimates of variable communalities. The use of squared multiple correlations has been suggested when importance is placed on the latent constructs defining a set of item because it results in factors representing only the common variance among observed variables. Factors were retained based on Kaiser's criterion (eigenvalue > 1.0), visual examination of scree plots (Cattel's test), and subjective interpretation. The direct oblimin method (delta = 0) was used to rotate factors. Factor loadings with values less than 0.30 were not considered significant for factor interpretation.

Confirmatory factor analysis. All confirmatory factor analysis (CFA) models were tested using Amos 5.0 with full-information maximum likelihood estimation (Arbuckle, 1996; Arbuckle & Wothke, 1999; Enders & Bandalos, 2001). A total of 2.8% of the item responses were missing. Full-information maximum likelihood estimation has been shown to yield accurate fit indices and parameter estimates with up to 25% simulated missing data and is considered one of the best

methods for handling missing data in CFA (Arbuckle, 1996; Arbuckle & Wothke, 1999; Enders & Bandalos, 2001). Results were also obtained using LISREL 8.5 (Jöreskog & Sörbom, 1996).

The root mean square error of approximation (RMSEA), non-normed fit index (NNFI), comparative fit index (CFI), expected cross-validation index (ECVI), and χ^2 statistic were used to evaluate and compare model fit. The χ^2 statistic is very sensitive to sample size and, in most CFA analyses, indicates rejection of the hypothesized model (Bollen, 1989; Joreskog, 1993; Long, 1983). For this reason, it is reported but supplemented by other fit indices to draw specific conclusions about model fit. The RMSEA is a standardized estimation that represents closeness of fit of population data to the model. It is widely used and is considered one of the most informative fit criteria (Byrne, 1998). Values approximating 0.06 to 0.08 reflect close and reasonable fit of the model (Browne & Cudeck, 1993; Hu & Bentler, 1998). The 90% confidence interval (CI) around the RMSEA point estimate should contain 0.06 or zero to indicate the possibility of close or exact fit. For all multiple groups analyses the RMSEA estimate and CI were multiplied by $\sqrt{2}$ as suggested by Steiger (1998). The CFI and NNFI test the proportionate improvement in fit by comparing the target model to some baseline model (Bentler & Bonett, 1980; Hu & Bentler, 1998). The NNFI, but not the CFI, is affected by model parsimony (more complex models are penalized). Values for the CFI and NNFI approximating 0.95 indicate good fit (Hu & Bentler, 1999). Acceptable model fit was based on CFI and NNFI values of 0.90 (Akaike, 1987; Bandalos, 1993; Bentler & Bonett, 1980; Browne & Cudeck, 1989; Hu & Bentler, 1998; Hu & Bentler, 1999). The ECVI was also used to compare models. The ECVI takes into account both model fit and parsimony and reflects how well a model might cross-validate in similar samples. A lower value for the ECVI indicates a model with fit and parameter values more likely to replicate in future samples (Akaike 1987; Browne and Cudeck 1989; Bandalos 1993; Hu and Bentler 1998). Overall judgment of model fit was based on model interpretability, the cumulative evidence from all measures of fit, and current suggestions that strict

"cut-off" values for specific indices may not always be appropriate(Byrne, 1998; Hu & Bentler, 1999; Marsh, Hau, & Wen, 2004).

In addition to the individual CFA models, a final model including the PACES-M and PACES was estimated. The measurement structure of each instrument was the same as specified in Figures 1 and 2. The purpose of this model was to examine the relationships among the factors represented by the two scales. In addition to the measurement portion of each model, the correlations among the three factors were included.

Invariance. Multi-group factor invariance was established by testing and comparing a series of models using standard procedures (Horn & McArdle, 1992; Vandenberg & Lance, 2000). Measurement equivalence between groups should be established to support conclusions and implications drawn from instruments designed to tap latent constructs. The models shown in Figures 1 and 2 were initially tested for each group (black, white, male, female) separately. This allowed the adequacy of the model to be assessed within each group prior to the multi-group invariance analysis. Due to the limited number of black participants (n = 192) invariance was not tested across four groups (i.e. black male, white males, black females, white females), rather two separate invariance analyses (one for race and one for gender) were conducted. Model specifications were the same for the gender and racial analyses.

The invariance analysis involved testing and comparing a series of nested models. In confirmatory factor analysis, models are considered nested if a simpler model can be obtained by imposing a set of restrictions on a more complex model. The only difference between the models is the number of free and fixed parameters estimated (i.e. the nested, or simpler, model has more fixed, or constrained parameters). Because the factor structure for the PACES (two correlated factors) was different than the PACES-M (1-factor plus method effect) the parameter constraints imposed on successive models are slightly different. For this reason, the invariance models tested are presented

separately for PACES and PACES-M. Throughout the text and tables, models for the PACES include the subscript O (original) while PACES-M models include the subscript M (modified).

For the PACES, four nested models were tested. Each model ($M1_O$ to $M4_O$) included previous model restrictions (i.e. $M3_O$ included restrictions from $M2_O$) plus additional constraints. In model 1 ($M1_O$), all hypothesized parameters were freely estimated in the two groups (males and females; black and white). This allowed the equivalence of the hypothesized pattern of paths, factor variance and covariance, and item uniquenesses to be evaluated. In model 2 ($M2_O$) the paths from the enjoyment and energy factors to the observed items were restricted to be equal across groups. The equivalence of models 1 and 2 was considered sufficient evidence to conclude that the measurement structure of the scales was invariant. Additional constraints were tested as follows. In model 3 ($M3_O$), factor variance and covariance estimates were added to those being held invariant. This included restricting two factor variances and a single covariance. Finally, in model 4 ($M4_O$) the observed item uniquenesses (measurement error term) were held equivalent across groups. Item uniquenesses reflect both random and systematic variance not explained by the factor. Testing their equivalence is very restrictive and does not need to hold for an instrument to be considered invariant across groups (Byrne, 1998)

For the PACES-M five nested models were tested. The constraints imposed for Model 1 ($M1_M$) were used to examine the equivalence of the hypothesized pattern of paths, factor variance and covariance, and item uniquenesses across groups. Model 2 ($M2_M$) was specified to restrict paths from the enjoyment factor to the observed items to be equal across groups. In model 3 ($M3_M$), the enjoyment factor variance was held invariant. Model 4 ($M4_M$) was specified so the correlated uniqueness representing the item method effect were constrained to be equal. Finally, in model 5 ($M5_M$) each observed item uniqueness was held equivalent across groups.

When fit indices are compared across nested models it can be determined if model fit is affected by constraining parameters to be equal across groups. Equivalence of factor structure and factor loadings was considered the minimal criteria for concluding invariance across group. Although χ^2 difference tests were conducted and reported, the χ^2 difference test is sensitive to sample size and thus conclusions for the invariance analysis were based primarily on changes in the values of the RMSEA, NNFI, and CFI between nested models. An absolute change in CFI of 0.01 between nested models has been reported to work well for testing multi-group invariance (Cheung & Rensvold, 2002). Subjective assessment of RMSEA point estimates and 90% CIs between two nested models were also used to judge meaningful change in fit between models. Based on current recommendations, parameter estimates, standard errors, z-values, and squared multiple correlations were inspected for sign and magnitude (Boomsma, 2000; Joreskog, 1993).

Results

Univariate and multivariate skewness and kurtosis values were obtained using PRELIS 2.5 (Joreskog & Sorbom, 1996). Mardia's coefficient of multivariate kurtosis was significant for the PACES ($z = 36.89$, $p < 0.001$) and the PACES-M ($z = 34.33$, $p < 0.001$), indicating a violation of multivariate normality. This coefficient is, however, sensitive to large sample sizes and is often significant when model fit, parameter estimates, and standard errors would not be substantially affected (Bollen, 1989; Kline, 1998; Mardia, 1970). Examination of the relative multivariate kurtosis index and univariate kurtosis values indicated that violations of multivariate normality were minimal and not likely to substantially affect the model estimates (Kline, 1998). For the PACES, absolute kurtosis ranged from 0.004 to 2.44 (mean of absolute kurtosis values = 0.813; relative multivariate kurtosis = 1.51). Only one item (depressed…happy) had a value greater than 2.0. For the PACES-M, absolute kurtosis ranged from 0.026 to 4.97 (mean of absolute kurtosis values = 1.25; relative

45

multivariate kurtosis = 1.51). One item (depressed) had a value greater than 4.0, while 3 others had values between 2.0 and 2.57.

Mean, standard deviation, and sample size for each item can be found in Tables 3.1 and 3.2. The lowest item means for both scales, across all groups was for "rather be doing something else". For females, the highest item means were for the "depressed" item in both scales. In the PACES-M, males also rated this item highest, but for the PACES they rated the "enjoy…hate" item higher. The internal consistency reliability coefficients were very similar for both scales across race and sex, ranging from 0.93 to 0.94. Correlations with the social desirability scale were small for both the PACES (r = 0.089) and PACES-M (r = 0.151).

Factor validity and invariance

PACES-M. The hypothesized single factor model with a method effect for positively worded items (see figure 3.1) provided a good fit to the data (CFI=0.977, NNFI=0.953, RMSEA [90% CI] = 0.055 [0.048 – 0.061]). All parameter estimates and correlated uniquenesses were significant and of the appropriate direction and magnitude. Squared multiple correlations for the items ranged from 0.185 to 0.666. The fit indices for three additional models estimated to examine the method effect previously reported for the PACES-M are provided in Table 3.3. The single factor model did not fit well. Although acceptable, the fit of the two factor model and the single factor correlated uniqunesses model (negative) were worse than the hypothesized model. Standardized parameter estimates and squared multiple correlations for the hypothesized model are presented in Table 3.2.

The fit of the hypothesized model for the PACES-M was good for racial and gender groups. The values of CFI (~ 0.98), NNFI (~0.95), and RMSEA (~0.055) were nearly identical between black and white and male and female participants. The χ^2 difference test and change in CFI, NNFI, and RMSEA indicate that factor loadings and the factor variance are equivalent between black and

46

white participants and between male and female participants. There was also very little change in CFI, NNFI, and RMSEA when correlated errors among positive items and item uniquenesses were held invariant.

PACES. The hypothesized single factor model fit poorly in the entire sample (CFI=0.844, NNFI=0.803, RMSEA [90% CI] = 0.112 [0.107 – 0.116]). Exploratory factor analysis conducted with data from the test sample suggested a two factor solution. In addition, items 7, 10, and 15 had loadings between 0.30 and 0.40 on one or both factors. The first factor, PACES-enjoy, accounted for 48.1% of the variance and included items reflecting enjoyment, interest, fun, and liking. The other factor, PACES-energy, included five items related to energy and achievement, accounting for 5.8% of the variance among items. Based on these findings and a previous confirmatory analysis of the PACES (Crocker et al., 1995), a two factor model which included 15 (items 7, 10, and 15 were dropped) of the original 18 items was specified (see Figure 3.2).

Results of the CFAs for the two factor model, a single factor model including all 18 items, and an alternate single factor model including the 15 items retained from the exploratory analysis are presented in Table 3.4. The 15- and 18-item single factor models did not fit well in the test or validation samples. The two factor model developed with the test sample cross-validated quite well. The fit indices indicate acceptable to good fit in both the test and validation samples. Standardized parameter estimates and squared multiple correlations for this model are presented in Table 3.1.

The two factor, 15-item model fit well in the entire sample (CFI=0.946, NNFI=0.927, RMSEA [90% CI] = 0.075 [0.070 – 0.081]) and is adequate for racial and gender groups (see Table 3.5). The fit for the sample of black participants was, however, questionable. Because all measures of fit presented are affected by sample size, two post-hoc analyses were conducted to determine if this difference in fit was a function of sample size or race. First, an exploratory factor analysis was conducted for black and white participants separately. A very similar two factor solution emerged

47

for both samples. Second, a random sample of white participants (n = 192) equaling the number of black participants was created. The model (see Figure 3.2) was then fit to data from this sample of white individuals. Model fit (χ^2 = 228.51, CFI=0.920, NNFI=0.892, RMSEA [90% CI] = 0.091 [0.076 – 0.105]) was worse than for the entire white sample, but slightly better than for the black sample. The similarity in model fit between the white and black samples of similar size and the change in fit that occurred with a smaller sample of white participants seemed to indicate that sample size accounted for most of the difference in fit between groups. For this reason the two-factor (15-item) model was retained for both racial groups.

The invariance of the two-factor model between white and black participants was supported see Table 3.5). The χ^2 difference test and change in CFI, NNFI, and RMSEA indicate that factor loadings and the factor variance were equivalent between black and white participants. NNFI, CFI, and RMSEA also change very little when factor variances and the factor covariance were held invariant. Evidence for the equivalence of item uniquenesses between black and white participants is marginal. The χ^2 difference test was significant and the change in CFI was very close to -0.01, however, RMSEA changed very little.

The invariance of the two-factor model between male and female participants was supported see Table 3.5). Model fit changed very little for $M1_O$, $M2_O$, and $M3_O$. The χ^2 difference tests were significant, but changes in CFI, NNFI and RMSEA were minimal. There was also very little change in CFI, NNFI, and RMSEA when item uniquenesses were held invariant ($M4_O$).

Correlations among the scale total scores (PACES-M, PACES-enjoy, PACES-energy) can be found in Table 3.6. Bivariate correlations were similar between racial and gender groups. Total scores from PACES-M and PACES-enjoy correlated highly with each other in the total sample (r = 0.721). The fit of the model which included both scales was acceptable (CFI=0.927, NNFI=0.908, RMSEA [90% CI] = 0.060 [0.057 – 0.063]). The factor correlations among PACES-M, PACES-

enjoy, and PACES-energy were similar in magnitude to the bivariate correlations. The factor

correlations between PACES-M and the two PACES factors were 0.764 and 0.620 for PACES-enjoy

and PACES-energy. The factor correlation between PACES-enjoy and PACES-energy was 0.748.

The difference in PACES-M total score between white and black participants was not

statistically significant (t = 1.094, p = .274). There was a significant difference between males and

females (t = 2.314, p = 0.02). Because the factor structure of the PACES has been brought into

question, t-tests using PACES total score (all 18-items), PACES-energy score (5-items), and PACES-

enjoy score (10-items) were conducted. Significant differences were found for PACES total score

between groups. Males' scores were higher than females' (t = 4.524, p < .001), and scores for white

participants were higher compared to black participants (t = 2.496, p = 0.013). When scores were

calculated for the new two factor model of the PACES a slightly different pattern emerged. Again

males had higher scores than females for both PACES-energy (t = 2.647, p = 0.008) and PACES-

enjoy (t = 5.733, p < 0.001). In the racial comparison, a significant difference in PACES-energy was

found (t = 4.611, p < 0.001), but PACES-enjoy did not differ between black and white participants

(t = 1.484, p = 0.138).

Discussion

The purpose of this study was to: (1) evaluated the factor structure of the original PACES,

(2) cross validate the factor structure of the PACES-M, (3) compare the PACES and PACES-M, and

(4) test the factor invariance of the two scales across gender and race in a sample of young adults.

The results supported the hypothesized factor structure of the PACES-M. A single factor model did

not fit well for the PACES, but a two factor model did provide adequate fit to the data. In addition,

the factor invariance between black and white and male and female participants was supported for

both scales. The correlations among scale scores were moderate to high, supporting convergence of measurement.

The findings did not support the hypothesized single factor model for the PACES. These results are in agreement with the only other published confirmatory analysis of the PACES (Crocker et al., 1995). In the current investigation, the previous confirmatory work was extended by utilizing exploratory analyses to determine a better fitting model. These results were then cross-validated and confirmed in a separate sample of young adults. The two-factor model increases the complexity of interpreting scores from the PACES, but is a much better representation of the covariance among items on the scale. It was found that items reflecting "energy and achievement" loaded on one factor (PACES-energy), while items reflecting general positive affect loaded on the other (PACES-enjoy). These groupings are in partial agreement with observations made by Crocker et al. (1995). In their discussion they point out that PACES items may be tapping constructs such as movement sensation, competence motives, and positive affective states. Applying their example to these results, it seems that items indicating energy and achievement are those related to movement sensations and competence motives, while items indicating the PACES-enjoy factor reflect general positive affect. The results suggest that young adults make some distinction between general positive affective states and feelings of achievement and vitality associated with physical activity. Whether these factors represent two different constructs or two facets of enjoyment is still not clear.

The results supporting the factor structure of the PACES-M are nearly identical to those previously reported in samples of adolescent girls and young children (Moore et al., 2004; Motl et al., 2001). The best fitting model of the PACES-M included a single enjoyment factor and correlated uniquenesses among the positively worded items. These correlations are intended to represent a method effect for positive item structure. Method effects are thought to exist when the "true" relationship among items (i.e. all measure enjoyment) is obscured by some other similarity among

50

items (i.e. positively or negatively structured, similar item wording, etc.). The method effect for the PACES-M, although tenable, must be qualified. In actuality the correlations among the uniqueness represent a multidimensional method effect among the positively worded items. In other words, these correlations represent the method effect for positive items, but may also include effects of item wording, item order, etc. More accurately stated, the best fitting model for the PACES-M (see figure 3.2) includes a single enjoyment factor representing common variance among all items and the correlations among the uniquenesses for the positive items. The correlated uniquenesses represent shared variance among the positively worded items not accounted for by the enjoyment factor. This includes, but is not limited to, a method effect for positive items structure.

The results of the exploratory and confirmatory factor analyses do not support the use of PACES total score. For this reason, separate scores for PACES-energy and PACES-enjoy were calculated. Correlations among PACES and PACES-M scores were moderate to large and did not differ between males and females or between black and white participants. The factor correlation between PACES-M and PACES-enjoy was significantly larger than between PACES-M and PACES-energy. These results provide some support for the convergence of measurement for the PACES-M and PACES-enjoy, but the relationship is not as strong as would be expected. The content of both scales is similar (eight of the 10 PACES-enjoy items correspond to 10 of the 16 PACES-M items), but only 58% of the variability is shared.

Mean scores for the PACES-M were very similar to those reported for adolescent girls (Dishman et al., 2005), but slightly higher than those calculated in a sample of third graders (Moore et al., 2004). Significant differences between black and white children were also reported for the third grade students, a finding not substantiated by our results. The fact that no racial differences were found in PACES-M score or PACES-enjoy score in this study could be a function of age or socioeconomic status. Interestingly, PACES-energy scores were significantly different between races,

with black participants scoring lower on the energy scale compared to whites. This distinction is masked when total PACES score is used for comparison.

Significant gender differences were found for all PACES and PACES-M scores. In this sample, men consistently scored higher than women. The difference was more pronounced in scores from the PACES. Very similar mean scores and gender differences were reported for 338 university students participating in a physical activity intervention (Calfas et al., 1999), but no differences were found for 280 youth sports participants aged 12 to 16 (Crocker et al., 1995). The difference between the participants in this study and the adolescents could be age, but is more likely a result of activity choice and participation. It seems reasonable that girls who choose to attend sports camp will report higher levels of enjoyment of physical activity/sport compared to girls, or women, in the general population.

Because PACES total score is currently being used and reported it seems important to note that if scores from the PACES and PACES-M were used to compare black and white participants different conclusions would result. For the PACES, the difference in mean score was statistically significant, for the PACES-M there was not a significant difference. The separation of PACES total score into PACES-enjoy and PACES-energy sheds some light on this incongruence. The difference in PACES total score across groups seems to be a function of the energy items rather than the enjoyment items.

The invariance analysis provides additional support for the comparability of PACES and PACES-M scores between black and white and between male and female young adults. For both scales the factor structure and loadings were found to be equivalent between groups. In addition, the PACES-M models with restrictions placed on the factor variance, correlations among item uniquenesses, and item uniquenesses fit the data well with very little change in the CFI, RMSEA, or NNFI across models. This study provides the first evidence supporting the factorial invariance of

the PACES and PACES-M in young adults. Evidence from Motl et al (2001) suggests invariance between black and white girls, but provides no test of the PACES-M alone. Establishing measurement invariance is an important piece of validity evidence supporting the use and comparison of scores between groups. Although it does not rule out item bias, it can be concluded that the relationship among items on the PACES and PACES-M are nearly equivalent between male and female and between black and white young adults.

There is some indication that certain items within each scale need to be reevaluated or dropped. From the results of the exploratory factor analysis three PACES items with low factor loadings (0.30 to 0.40) and cross-loadings were identified. The items "energized", "good physically", and "stimulating" were subsequently dropped from the scale. One additional item, "Frustrated", was included in the confirmatory analysis, but was found to have a small squared multiple correlation. Two of these items, "stimulated" and "frustrated", were also identified by Crocker et al. (1995) as problematic. Although no PACES-M items were eliminated, six had squared multiple correlations below 0.30. Three of these items correspond to problem items from the PACES. The item on both scales which includes the term "depressed" may also need to be reassessed due to its moderate to large kurtosis and skewness values.

The results of this study provide additional evidence for the validity of interpretations for PACES and PACES-M scores. Although invariance was supported across race, generalization may be limited due to the small number of black males (n = 38) in the study. Attempts should be made to cross-validate these results with larger samples of black males and to extend these findings to other racial, ethnic, and cultural groups. The new two-factor model of the PACES, also, raises issue with the use of a total score to represent a single enjoyment construct. This model suggests that either enjoyment is a multidimensional construct or that the PACES measures two related, but

distinct latent factors.   The new PACES model was cross validated, but admittedly requires confirmation by other researchers in more diverse samples.

The PACES and PACES-M have been used to evaluate enjoyment in a variety of situations (aerobics, sports, physical activity, sedentary pursuits, etc.). To date, no direct comparisons have been made across activity modes. Changing the context of either scale (i.e. sports, fitness activities, vigorous activities, sedentary activities, etc.) may affect the measurement structure of the instrument. This is suggested as a next step in extending the evidence for these two measures. The evidence also suggests a need to reevaluate and possibly eliminate items that do not function well. If the construct can be measured just as well with 10-items it would be much more economical to use a shorter scale. If the validity evidence continues to support the use of scores from these scales over a large age range in diverse samples of males and females participating in various activities, these measures may prove useful in examining longitudinal change, comparing enjoyment across groups, and examining differences in enjoyment across activity modes.

References

Ajzen, I. (1988). <u>Attitudes, personality and behavior.</u> Chicago: Dorsey.

Ajzen, I., & Fishbein, M. (1974). Factors influencing intentions and the intention-behavior relation.
    <u>Human Relations, 27</u>(1), 1-15.

Akaike, H. (1987). Factor analysis and AIC. <u>Psychometrica, 52</u>, 317-332.

Arbuckle, J. L. (1996). Full information estimation in the presence of incomplete data. In G. A.
    Marcoulides & R. E. Schumacker (Eds.), <u>Advanced structural equation modeling: Issues and</u>
    <u>techniques</u> (pp. 243-277). Mahwah, NJ: Lawrence Erlbaum.

Arbuckle, J. L., & Wothke, W. (1999). AMOS 4.0 user's guide. Chicago, Il: SmallWaters Corporation.

Bandalos, D. L. (1993). Factors influencing cross-validation of confirmatory factor-analysis models.
    <u>Multivariate Behavioral Research, 28</u>(3), 351-374.

Bandura, A. (1986). <u>Social foundations of thought and action</u>. Englewood Cliffs, NJ: Prentice-Hall.

Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of
    covariance-structures. <u>Psychological Bulletin, 88</u>(3), 588-606.

Bollen, K. A. (1989). <u>Structural equations with latent variables.</u> New York: John Wiley.

Boomsma, A. (2000). Reporting analyses of covariance structures. <u>Structural Equation Modeling,</u>
    <u>7</u>(3), 461-483.

Browne, M. W., & Cudeck, R. (1989). Single sample cross-validation indexes for covariance-
    structures. <u>Multivariate Behavioral Research, 24</u>(4), 445-455.

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S.
    Long (Eds.), <u>Testing structural equation models</u> (pp. 136-162). Newbury Park, CA: Sage.

Byrne, B. M. (1998). <u>Structural equation modeling with LISREL, PRELIS, and SIMPLIS: Basic</u>
    <u>concepts, applications, and programming</u>. Mahwah, N.J.: Lawrence Erlbaum.

Calfas, K. J., Sallis, J. F., Nichols, J. F., Sarkin, J. A., Johnson, M. F., Caparosa, S., (2000). Project

GRAD: Two-year outcomes of a randomized controlled physical activity intervention among

young adults. American Journal of Preventive Medicine, 18(1), 28-37.

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing

measurement invariance. Structural Equation Modeling, 9(2), 233-255.

Crocker, P. R. E., Bouffard, M., & Gessaroli, M. E. (1995). Measuring enjoyment in youth sport

settings - A confirmatory factor-analysis of the physical-activity enjoyment scale. Journal of

Sport and Exercise Psychology, 17(2), 200-205.

Csikszentmihalyi, C. (1975). Beyond boredom and anxiety: Experience flow in work and play. San

Francisco: Jossey-Bass.

Deci, E. L. (1975). Intrinsic motivation. New York: Plenum.

Deci, E. L., & Ryan, R. M. (1985). Intrinsic motivation and self-determination in human behavior.

New York: Plenum.

DiLorenzo, T. M., Stucky-Ropp, R. C., Vander Wal, J. S., & Gotham, H. J. (1998). Determinants of

exercise among children. II. A longitudinal analysis. Preventive Medicine, 27(3), 470-477.

Dishman, R. K., Motl, R. W., Saunders, R., Felton, G., Ward, D. S., Dowda, M., et al. (2005).

Enjoyment mediates effects of a school-based physical-activity intervention. Medicine and

Science in Sports and Exercise, 37(3), 478-487.

Dishman, R. K., & Sallis, J. F. (1994). Determinants and interventions for physical activity and

exercise. In C. Bouchard, R. J. Shephard & T. Stephens (Eds.), Physical activity, fitness, and

health: International proceedings and consensus statement (pp. 214 - 238). Champaign, IL:

Human Kinetics.

Enders, C. K., & Bandalos, D. L. (2001). The relative performance of full information maximum likelihood estimation for missing data in structural equation models. Structural Equation Modeling, 8, 430-457.

Garcia, A. W., & King, A. C. (1991). Predicting long-term adherence to aerobic exercise: A comparison of two models. Journal of Sport and Exercise Psychology, 13, 394 - 410.

Horn, J. L., & McArdle, J. J. (1992). A practical and theoretical guide to measurement invariance in aging research. Experimental Aging Research, 18(3-4), 117-144.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. Structural Equation Modeling, 6(1).

Hu, L. T., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: sensitivity to underparameterized model misspecification. Psychological Methods, 3(4), 424-453.

Izard, C. E. (1977). Human emotions. New York: Plenum.

Johnson, N. A., & Heller, R. F. (1998). Prediction of patient nonadherence with home-based exercise for cardiac rehabilitation: The role of perceived barriers and perceived benefits. Preventive Medicine, 27(1), 56-64.

Joreskog, K. G. (1993). Testing structural equation models. In K. A. Bollen & J. S. Long (Eds.), Testing Structural Equation Models. Newbury Park, CA: Sage.

Jöreskog, K. G., & Sörbom, D. (1996a). LISREL 8: User's reference guide. Chicago, IL: Scientific Software International.

Joreskog, K., & Sorbom, D. (1996b). PRELIS 2: User's reference guide. Hillsdale, NJ: Lawrence Erlbaum.

Kendzierski, D., & Decarlo, K. J. (1991). Physical activity enjoyment scale - Two validation studies. Journal of Sport & Exercise Psychology, 13(1), 50-64.

Kimiecik, J. C., & Harris, A. T. (1996). What is enjoyment? A conceptual/definitional analysis with implications for sport and exercise psychology. Journal of Sport & Exercise Psychology, 18(3), 247-263.

Kline, R. B. (1998). Principles and practices of structural equation modeling. New York: Guilford.

Lewis, B. A., Marcus, B. H., Pate, R. R., & Dunn, A. L. (2002). Psychosocial mediators of physical activity behavior among adults and children. American Journal of Preventive Medicine, 23(2), 26-35.

Mardia, K. V. (1970). Measures of multivariate skewness and kurtosis with applications. Biometrika, 57, 519-530.

Marsh, H. W. (1996). Positive and negative global self-esteem: a substantively meaningful distinction or artifactors? Journal of Personality and Social Psychology, 70(4), 810-819.

Marsh, H. W., Hau, K., & Wen, Z. (2004). In search of golden rules: comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. Structural Equation Modeling, 11(3), 320-341.

Messick, S. (1995). Validity of psychological-assessment - validation of inferences from persons responses and performances as scientific inquiry into score meaning. American Psychologist, 50(9), 741-749.

Moore, J. B., Yin, Z., & Gutin, B. (2004). Enjoyment of physical activity in children: instrument validation in the MCG FitKid project. Journal of Sport and Exercise Psychology, 26, S138-S139.

Motl, R. W., Dishman, R. K., Saunders, R., Dowda, M., Felton, G., & Pate, R. R. (2001). Measuring enjoyment of physical activity in adolescent girls. American Journal of Preventive Medicine, 21(2), 110-117.

Rovniak, L. S., Anderson, E. S., Winett, R. A., & Stephens, R. S. (2002). Social cognitive determinants of physical activity in young adults: a prospective structural equation analysis. Annals of Behavioral Medicine, 24(2), 149-156.

Ryan, R. M., Frederick, C. M., Lepes, D., Rubio, N., & Sheldon, K. M. (1997). Intrinsic motivation and exercise adherence. International Journal of Sport Psychology, 28(4), 335-354.

Saelens, B. E., & Epstein, L. H. (1999). The rate of sedentary activities determines the reinforcing value of physical activity. Health Psychology, 18(6), 655-659.

Sallis, J. F. (1994). Influences on physical activity of children, adolescents, and adults. President's Council on Physical Fitness and Sport Research Digest, 1(7).

Sallis, J. F., Prochaska, J. J., & Taylor, W. C. (2000). A review of correlates of physical activity of children and adolescents. Medicine and Science in Sport and Exercise, 32, 963 - 975.

Salmon, J., Owen, N., Crawford, D., Bauman, A., & Sallis, J. F. (2003). Physical activity and sedentary behavior: A population-based study of barriers, enjoyment, and preference. Health Psychology, 22(2), 178-188.

Scanlan, T. K., Carpenter, P. J., Schmidt, G. W., Simons, J. P., & Keeler, B. (1993). An introduction to the sport commitment model. Journal of Sport and Exercise Psychology, 15(1), 1-15.

Scanlan, T. K., & Simons, J. P. (1992). The construct of sport enjoyment. In G. C. Roberts (Ed.), Motivation in sport and exercise (pp. 199-215). Champaign,IL: Human Kinetics.

Schleppe, S. (1993). The physical activity enjoyment scale: reliability and validity. University of Alberta (Canada), Canada.

Steiger, J. H. (1998). A note on multiple sample extensions of the RMSEA fit index. Structural Equation Modeling, 5(4), 411-419.

Stucky-Ropp, R. C., & DiLoenzo, T. M. (1993). Determinants of exercise in children. Preventive Medicine, 22, 880-889.

Thuot, S. M. (1995). College students' attitudes toward anticipated sport participation. Perceptual and Motor Skills, 80, 155-160.

US Department of Health and Human Services (1996). Physical activity and health: A report of the surgeon general. Atlanta, GA: Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion.

US Department of Health and Human Services (2003). Behavior risk factor surveillance system. Atlanta, GA:, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion.

US Department of Health and Human Services (2003). Youth risk behavior survey. Atlanta, GA:, Centers for Disease Control and Prevention, Chronic Disease and Health Promotion.

U.S. Department of Health and Human Services (2004). National health interview survey. Washington, D.C.: Public Health Service, Centers for Disease Control and Prevention, National Center for Health Statistics.

Vallerand, R. J. (2001). A hierarchical model of intrinsic and extrinsic motivation in sport and exercise. In G. C. Roberts (Ed.), Advances in motivation in sport and exercise. Champaign: Human Kinetics.

Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: suggestions, practices, and recommendations for organizational research. Organizational Research Methods, 3(1), 4-70.

Wankel, L. M. (1993). The importance of enjoyment to adherence and psychological benefits from physical-activity. International Journal of Sport Psychology, 24(2), 151-169.

Welk, G. J. (1999). The youth physical activity promotion model: a conceptual bridge between theory and practice. Quest, 51, 5-23.

Wininger, S. R. (2002). Instructors' and classroom characteristics associated with exercise enjoyment

    by females. <u>Perceptual and Motor Skills, 94</u>(2), 395-398.

Table 3.1 Sample size, means, standard deviations, and parameter estimates for items from the Physical Activity Enjoyment Scale (PACES).

#	PACES item			N	Total Mean	Total SD	Male Mean	Male SD	Female Mean	Female SD	White Mean	White SD	Black Mean	Black SD	Parameter Estimates Loading	Parameter Estimates SMC
1R	enjoy	…	hate	1003	5.91	1.28	6.23	1.03	5.76	1.36	6.00	1.20	5.64	1.41	0.839	0.704
2	bored	…	interested	1001	5.28	1.56	5.66	1.47	5.10	1.57	5.31	1.54	5.29	1.52	0.789	0.622
3	dislike	…	like	999	5.66	1.42	6.00	1.20	5.50	1.49	5.73	1.36	5.44	1.59	0.871	0.759
4R	pleasurable	…	unpleasurable	1000	5.43	1.52	5.71	1.40	5.30	1.55	5.51	1.45	5.24	1.68	0.766	0.587
5R	absorbed	…	not absorbed	1001	4.90	1.62	5.30	1.53	4.72	1.62	4.93	1.57	4.81	1.79	0.716	0.513
6	no fun	…	fun	998	5.43	1.40	5.78	1.33	5.27	1.41	5.46	1.36	5.38	1.50	0.822	0.675
7R	energizing	…	tiring	997	5.42	1.56	5.47	1.48	5.39	1.60	5.51	1.50	5.06	1.77	--	--
8	depressed	…	happy	993	5.93	1.20	6.02	1.07	5.88	1.26	6.00	1.16	5.72	1.30	0.674	0.454
9R	pleasant	…	unpleasant	998	5.48	1.41	5.78	1.26	5.34	1.45	5.54	1.35	5.29	1.60	0.748	0.559
10R	good physically	…	bad physically	996	5.82	1.33	5.94	1.19	5.77	1.39	5.91	1.28	5.59	1.45	--	--
11R	invigorating	…	not invigorating	1008	5.52	1.38	5.70	1.34	5.43	1.39	5.62	1.31	5.16	1.55	**0.744**	0.553
12	frustrated	…	not frustrated	1005	5.29	1.54	5.41	1.51	5.24	1.55	5.30	1.51	5.32	1.65	0.490	0.240
13R	gratifying	…	not gratifying	1016	5.70	1.37	5.84	1.24	5.64	1.42	5.82	1.28	5.28	1.57	**0.841**	0.708
14R	exhilarating	…	not exhilarating	1003	5.41	1.39	5.70	1.27	5.28	1.43	5.48	1.34	5.21	1.51	**0.839**	0.704
15	not stimulating	…	stimulating	1003	5.46	1.46	5.63	1.44	5.38	1.46	5.51	1.39	5.35	1.60	--	--
16R	strong sense of accomplishment	…	not sense of accomplishment	1016	5.70	1.50	5.68	1.49	5.71	1.51	5.83	1.37	5.32	1.82	**0.700**	0.489
17R	refreshing	…	not refreshing	1001	5.79	1.36	5.84	1.25	5.77	1.41	5.90	1.24	5.48	1.63	**0.761**	0.580
18	rather be doing something else	…	nothing else rather be doing	1005	4.72	1.58	5.06	1.54	4.56	1.57	4.70	1.57	4.78	1.59	0.702	0.493
			Total Score	950	98.89	18.10	102.84	16.47	97.14	18.53	99.93	17.41	96.10	19.66		
			PACES-Energy	996	28.15	5.76	28.86	5.21	27.83	5.96	28.66	5.28	26.52	6.86		
			PACES-Enjoy	969	59.41	12	62.69	10.8	57.93	12.3	59.91	11.70	58.25	12.9		

Note: The "R" next to the item number indicates the item has been reversed prior to analysis. Sample sizes used to calculate means and standard deviations varied from item to items: Male n= 302-322; Female n= 671-694; White n= 711-750; Black n= 174 – 192. The parameter estimates are from the two-factor (15-item) model. Loadings in **bold** are for the PACES-Energy factor. Those not in bold are for the PACES-Enjoy. Loading = standardized regression weight; SMC = squared multiple correlation; SD = standard deviation.

Table 3.2. Sample size, mean, standard deviation, and parameter estimates for items from the Modified Physical Activity Enjoyment Scale (PACES-M).

#	PACES-M item	N	Total Mean	Total SD	Male Mean	Male SD	Female Mean	Female SD	White Mean	White SD	Black Mean	Black SD	Parameter Estimates Loading	Parameter Estimates SMC
1	I enjoy it.	988	4.31	0.87	4.40	0.76	4.26	0.91	4.37	0.81	4.09	1.01	0.724	0.525
2R	I feel bored.	987	3.95	1.05	4.07	1.05	3.90	1.05	3.93	1.05	4.10	1.02	0.735	0.540
3R	I dislike it.	985	4.06	1.04	4.19	0.96	3.99	1.07	4.08	1.02	4.02	1.07	0.786	0.618
4	I get something out of it.	981	4.42	0.75	4.40	0.70	4.42	0.77	4.46	0.72	4.34	0.78	0.504	0.254
5R	It's no fun at all.	987	4.23	0.94	4.33	0.89	4.18	0.96	4.25	0.91	4.18	0.99	0.816	0.666
6	It gives me energy.	986	4.28	0.81	4.18	0.81	4.33	0.80	4.35	0.74	4.10	0.93	0.471	0.222
7R	It makes me depressed.	986	4.60	0.76	4.60	0.76	4.60	0.76	4.61	0.74	4.61	0.76	0.533	0.285
8	It's very pleasant.	986	3.92	0.93	4.03	0.86	3.87	0.96	3.96	0.90	3.87	0.98	0.698	0.488
9	My body feels good.	987	4.30	0.85	4.25	0.86	4.32	0.84	4.36	0.79	4.07	0.97	0.535	0.286
10	I find it fun.	989	3.99	0.96	4.20	0.86	3.89	0.99	4.02	0.92	3.93	1.01	0.738	0.545
11	It's very exciting.	985	3.94	0.97	4.12	0.90	3.85	0.99	3.94	0.95	3.92	1.02	0.691	0.477
12R	It frustrates me.	984	3.89	1.06	3.96	0.99	3.86	1.10	3.87	1.06	3.96	1.10	0.542	0.294
13R	It's not at all interesting.	981	4.03	1.02	4.12	1.03	4.00	1.02	4.01	1.02	4.22	0.98	0.687	0.472
14	It gives me a strong feeling of success.	987	4.20	0.86	4.13	0.83	4.23	0.87	4.24	0.83	4.05	0.94	0.431	0.185
15	It feels good.	979	4.32	0.82	4.37	0.80	4.30	0.83	4.37	0.79	4.14	0.89	0.629	0.396
16R	I feel as though I would rather be doing something else.	985	3.73	1.13	3.88	1.10	3.65	1.13	3.70	1.12	3.80	1.13	0.735	0.540
	Total Score	929	66.18	10.32	67.32	9.71	65.64	10.56	66.51	10.03	65.55	10.99		

Note: The "R" next to the item number indicates the item has been reversed prior to analysis. Sample sizes used to calculate means and standard deviations varied from item to items: Male n= 299-317; Female n= 630-672; White n= 692-731; Black n= 173 – 189. The parameter estimates are from the single factor model with correlated uniquenesses among the positive items. Loading = standardized regression weights; SMC = squared multiple correlation; SD = standard deviation.

Table 3.3. Measures of fit and model comparisons for the PACES-M.

	χ^2	DF	NNFI	CFI	ECVI	RMSEA (90% CI)	Models	$\Delta\chi^2$	Δ df	p	Δ CFI
Factor Structure											
Single Factor - CU(positive)*	275.2	68	0.953	0.977	0.434	0.055 (0.048 - 0.061)					
Single Factor - CU(negative)	844.02	83	0.859	0.914	0.961	0.095 (0.089 - 0.101)					
Two Method Factors (pos/neg)	924.14	103	0.877	0.907	1.000	0.088 (0.083 - 0.094)					
Single Factor	1325.48	104	0.819	0.862	1.391	0.107 (0.102 - 0.112)					
							Model Comparisons				
Tests of Invariance - Race							**Models**	$\Delta\chi^2$	Δ df	p	Δ CFI
Black	107.435	68	0.956	0.978	1.444	0.055 (0.034 - 0.074)					
White	224.396	68	0.951	0.976	0.52	0.055 (0.047 - 0.063)					
M1$_M$ - factor structure	331.989	136	0.952	0.976	0.706	0.055 (0.048 - 0.062)					
M2$_M$ - factor loadings	345.394	151	0.958	0.976	0.689	0.052 (0.045 − 0.054)	M1 v M2	13.41	15	0.571	0.000
M3$_M$ - factor variance	346.526	152	0.958	0.976	0.688	0.052 (0.045 − 0.059)	M2 v M3	1.132	1	0.287	0.000
M4$_M$ - CU representing method effect	418.116	188	0.960	0.972	0.687	0.051 (0.044 − 0.058)	M3 v M4	71.59	36	< 0.001	-0.004
M5$_M$ - item uniquenesses	475.313	204	0.956	0.967	0.714	0.052 (0.045 − 0.059)	M4 v M5	57.2	16	< 0.001	-0.005
Tests of Invariance - Gender											
Male	120.022	68	0.961	0.980	0.884	0.048 (0.034 - 0.062)					
Female	235.82	68	0.946	0.973	0.581	0.060 (0.051 - 0.068)					
M1$_M$ - factor structure	355.852	136	0.950	0.975	0.678	0.057 (0.049 - 0.064)					
M2$_M$ - factor loadings	373.35	151	0.955	0.975	0.665	0.054 (0.047 - 0.061)	M1 v M2	17.5	15	0.290	0.000
M3$_M$ - factor variance	376.994	152	0.954	0.975	0.667	0.054 (0.047 - 0.061)	M2 v M3	3.644	1	0.056	0.000
M4$_M$ - CU representing method effect	450.178	188	0.957	0.970	0.668	0.052 (0.047 - 0.058)	M3 v M4	73.18	36	< 0.001	-0.005
M5$_M$ - item uniquenesses	484.979	204	0.958	0.968	0.671	0.052 (0.047 - 0.058)	M4 v M5	34.8	16	0.004	-0.002

Note: * denotes hypothesized model. CU = Correlated uniquenesses. The measures of fit include: χ^2 = Chi-square; DF = degrees of freedom; NNFI = Non-normed fit index; CFI = comparative fit index; ECVI= Expected cross-validation index; RMSEA = Root mean square error of approximation. For model comparisons Δ represents change in a given value between models.

Table 3.4. Fit for hypothesized and exploratory models of PACES in test and validation samples

Model	χ^2	DF	NNFI	CFI	ECVI	RMSEA (90% CI)
Total Sample (N=1023)						
Single Factor (18-items)	1858.54	135	0.803	0.844	1.924	0.112 (0.107 - 0.116)
Test Sample (n = 409)						
Single Factor (18-items)	903.01	135	0.782	0.828	2.379	0.116 (0.109 - 0.123)
Single factor (15-items)	737.765	90	0.773	0.830	1.948	0.130 (0.122 - 0.139)
Two factors (15-items; figure 2)	317.895	89	0.919	0.940	0.964	0.078 (0.069 - 0.087)
Validation Sample (n = 614)						
Single Factor (18-items)	1135.49	135	0.809	0.849	2.086	0.112 (0.106 - 0.118)
Single factor (15-items)	921.685	90	0.808	0.856	1.697	0.125 (0.117 - 0.132)
Two factors (15-items; figure 2)	414.189	89	0.924	0.944	0.849	0.078 (0.071 - 0.071)

Note: The measures of fit include: χ^2 = Chi-square; DF = degrees of freedom; NNFI = Non-normed fit index; CFI = comparative fit index; ECVI= Expected cross-validation index; RMSEA = Root mean square error of approximation.

Table 3.5. Measures of fit and model comparisons for the PACES invariance analysis.

	χ^2	DF	NNFI	CFI	ECVI	RMSEA (90% CI)					
Factor Structure											
Two factors (15-items; figure 2)	604.448	89	0.927	0.946	0.681	0.075 (0.070 - 0.081)					
							Model Comparisons				
Tests of Invariance - Race							Models	$\Delta \chi^2$	Δ df	p	Δ CFI
Black (n = 192)	284.829	89	0.866	0.901	1.973	0.107 (0.093 - 0.121)					
White (n = 756)	445.408	89	0.930	0.948	0.712	0.073 (0.066 - 0.080)					
$M1_O$ - factor structure	730.773	178	0.916	0.937	0.967	0.081 (0.075 - 0.088)					
$M2_O$ - factor loadings	744.75	191	0.921	0.937	0.954	0.078 (0.072 - 0.085)	M1 v M2	13.98	13	0.375	0.000
$M3_O$ - factor Var/Cov	764.408	194	0.920	0.935	0.969	0.079 (0.074 - 0.085)	M2 v M3	19.66	3	<0.001	-0.002
$M4_O$ - item uniquenesses	862.009	209	0.915	0.926	1.04	0.081 (0.076 - 0.085)	M3 v M4	97.6	15	<0.001	-0.009
Tests of Invariance - Gender											
Male (n = 327)	287.997	89	0.902	0.928	1.166	0.083 (0.072 - 0.094)					
Female (n = 696)	457.955	89	0.926	0.945	0.791	0.077 (0.070 - 0.084)					
$M1_O$ - factor structure	746.033	178	0.919	0.940	0.911	0.079 (0.074 - 0.085)					
$M2_O$ - factor loadings	803.797	191	0.919	0.935	0.942	0.079 (0.074 - 0.085)	M1 v M2	57.76	13	<0.001	-0.005
$M3_O$ - factor Var/Cov	819.654	194	0.918	0.934	0.952	0.079 (0.074 - 0.085)	M2 v M3	15.86	3	0.001	-0.001
$M4_O$ - item uniquenesses	893.778	209	0.917	0.928	0.995	0.081 (0.075 - 0.085)	M3 v M4	74.12	15	<0.001	-0.006

Note: The measures of fit include: χ^2 = Chi-square; DF = degrees of freedom; NNFI = Non-normed fit index; CFI = comparative fit index; ECVI= Expected cross-validation index; RMSEA = Root mean square error of approximation. For model comparisons Δ represents change in a given value between models.

Table 3.6. Bivariate correlations among PACES-M, PACES-Joy, and PACES-Energy total scores for male and female and black and white participants.

		Bivariate Correlations					
		Gender			Race		
		1	2	3	1	2	3
1	PACES-M		0.724	0.587		0.666	0.615
2	PACES-Joy	0.715		0.731	0.732		0.763
3	PACES-Energy	0.650	0.710		0.601	0.715	

Note: Correlations for Female and Black participants are above the diagonals for their respective group.
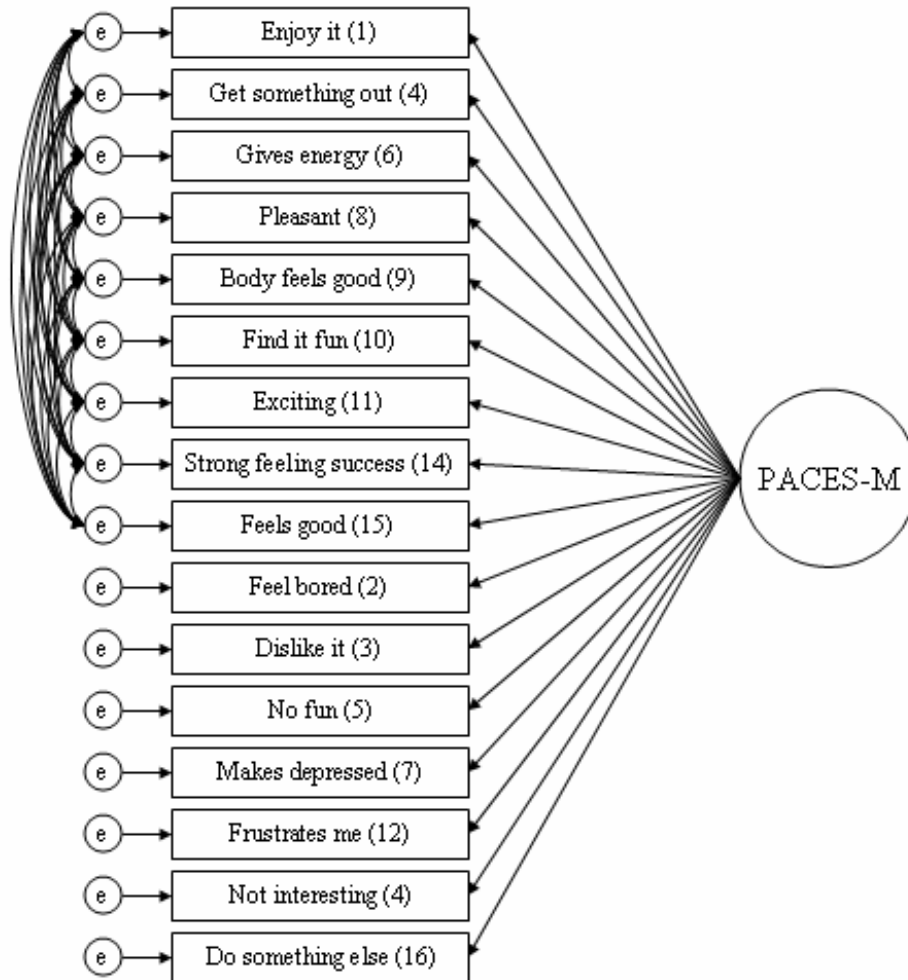
Figure 3.1. Measurement model for the Modified Physical Activity Enjoyment Scale (PACES-M). The model includes the 18 items (item#) indicating enjoyment of physical activity and correlations among the uniquenesses for the positively worded items. The "e", or item uniqueness, represents the random and systematic variance not accounted for by the factor PACES-M.
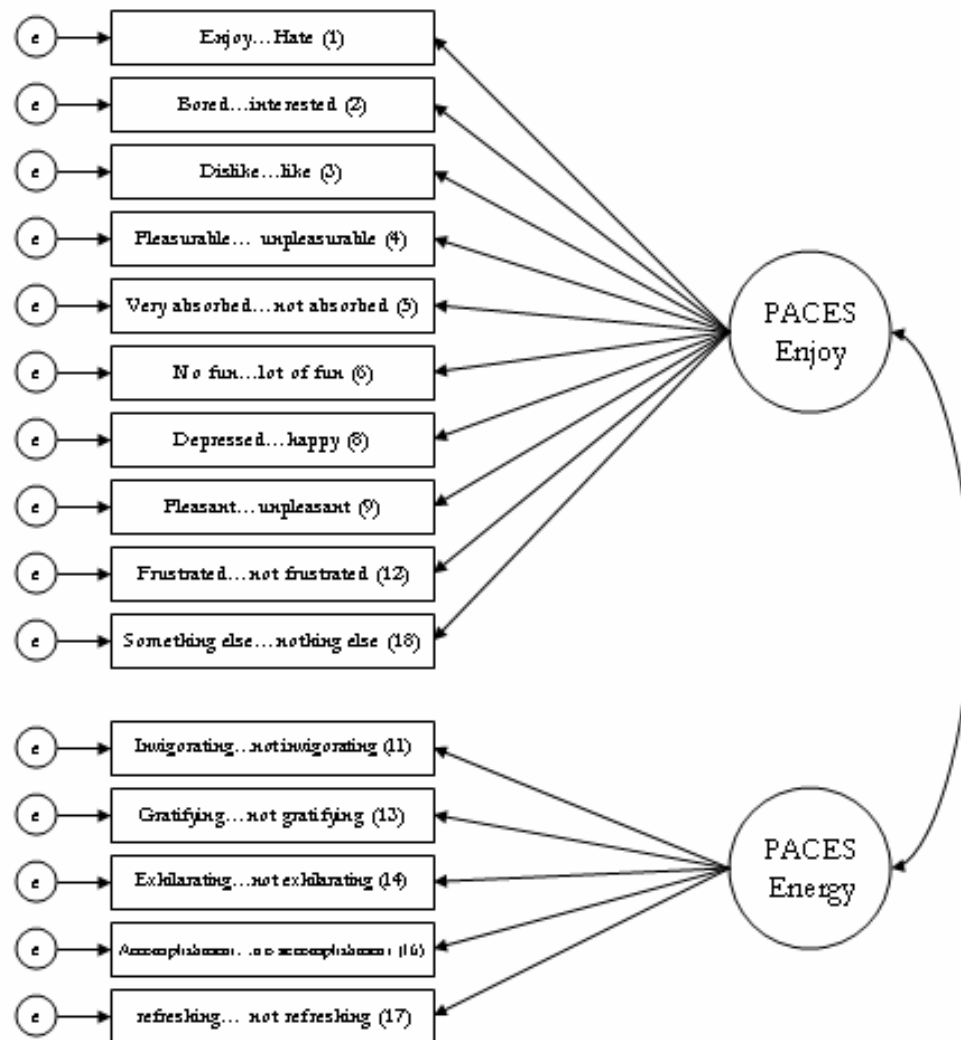
Figure 3.2. Measurement model for the Physical Activity Enjoyment Scale (PACES). The model includes two correlated factors, PACES-enjoy and PACES-energy, indicated by 10 and five items respectively. Item number is in parentheses. The "e", or item uniqueness, represents the random and systematic variance not accounted for by the factor for which the item is an indicator.

CHAPTER 4

CONSTRUCT VALIDITY AND COMPARISON OF SEVERAL MEASURES OF PHYSICAL

ACTIVITY ENJOYMENT[1]

Abstract

Enjoyment is often cited as an important correlate of physical activity behavior and is included in a number of health promotion models and theories of behavior motivation. A number of measures have been developed and used to quantify enjoyment of physical activity but most lack sufficient validity evidence. In addition, little is known about how the various measures interrelate. Therefore, the purpose of these studies was to compare and extend the construct validity evidence for several measures of physical activity enjoyment. Five measures were chosen for comparison. The convergence of measures was examined using confirmatory factor analysis. In addition, each measure of enjoyment was correlated with measures of physical activity, perceived competence and extrinsic motives for physical activity. Participants (N = 1023; mean age 19.60 ± 1.55 years) were recruited from basic physical education classes and the psychology research pool at two universities. The results of the confirmatory factor and correlational analyses supported the similarity of enjoyment scale scores, but suggested that measuring enjoyment and perceived competence with the same instrument inflates their relationship. Invariance analyses supported the comparability of scores between black and white and male and female young adults. Each measure of enjoyment was modestly correlated with total and vigorous physical activity (0.22 – 0.34), but minimally related to work and moderate activity (0.047 – 0.12). As expected, the correlations with perceived competence were larger than with the extrinsic motives. The magnitude of these relationships was very similar across the five measures of enjoyment. Overall, this research provides support for the construct validity of scores from several measures of physical activity enjoyment.

Key Words: Enjoyment; Confirmatory factor analysis; Validity; Physical Activity

Introduction

Enjoyment has been cited as an important indicator, or outcome, of intrinsically motivated behavior (Deci & Ryan, 1985). It is an inherent aspect of most freely chosen leisure activities (Wankel, 1993) and is considered one of the core emotions linked to intrinsically motivated human behavior (Izard, 1977). In the fields of leisure, health and exercise science, enjoyment is generally defined as a positive affective state either cognitive or physiological, involving feelings of pleasure, fun, and liking (Scanlan & Simons, 1992; Wankel, 1993). This definition has been debated (Kimiecik & Harris, 1996; Wankel, 1997), but items designed to operationalize the construct of enjoyment most often reflect the definitions presented by Scanlan and Simons (1992) and Wankel (1993). In fact, the descriptors "fun", "enjoy", "liking", and "interest/pleasure" are those used most often for items measuring enjoyment of exercise, sport, or physical activity.

Because important relationships between enjoyment of physical activity and physical activity behavior have been noted for both children and adults (DiLorenzo, Stucky-Ropp, Vander Wal, & Gotham, 1998; Dishman & Sallis, 1994; Garcia & King, 1991; Rovniak, Anderson, Winett, & Stephens, 2002; Sallis, 1994; Sallis, Prochaska, & Taylor, 2000; Salmon, Owen, Crawford, Bauman, & Sallis, 2003; Stucky-Ropp & DiLoenzo, 1993; Thuot, 1995) and because various theories and health promotion models include the construct of enjoyment, either explicitly or implicitly (Ajzen, 1988; Ajzen & Fishbein, 1974; Bandura, 1986; Deci & Ryan, 1985; Vallerand, 2001; Welk, 1999) a number of subscales, stand-alone scales, and single item scales have been used to measure enjoyment of exercise, physical activity, and/or sport. These scales/items are commonly used to investigate mediator-moderator relationships (Dishman et al., 2005; Motl et al., 2001), examine intervention outcomes (Calfas et al., 2000; Stevens, Lemmink, van Heuvelen, de Jong, & Rispens, 2003), and compare groups, both experimental and demographic (Moore, Yin, & Gutin, 2004; Ryan, Frederick, Lepes, Rubio, & Sheldon, 1997; Whitehead & Corbin, 1991). Despite this use, very little is

72

known about how the scores and psychometric properties of the different measures of enjoyment compare.

Only one published study was found in which a correlation between two enjoyment measures was reported (Ryska, 2003). The researchers noted a moderate correlation ($r = 0.461$), indicating that despite similar names and uses (measure enjoyment/interest) the scales may not be interchangeable. It is often assumed that tests with similar names measure the same construct, and that tests with different names measure different constructs. The current recommendations suggest that convergent and discriminant validity evidence be provided to support these assumptions (AERA, 1999; Messick, 1995). When examining multiple measures of the same construct, three important questions should be addressed during validation: (1) To what degree do the measures of the same construct relate?, (2) Is the construct in question distinct from other related factors?, and (3) Is the magnitude of the relationship between the construct of interest and the other related factors similar across all measures of the construct? Without support for convergence, discrimination, and comparability of external relationships, one or all measures must be called into question.

The result of having scales with similar names that measure different, or slightly different, constructs is that comparison of results becomes difficult, if not impossible. If scales intended to measure the same thing converge, results from different studies can be directly compared providing more robust estimates of theoretical relationships. Currently, no evidence for the convergent validity of measures of enjoyment is available. This fact, along with recent suggestions calling for comparisons of similarly named constructs used in different behavioral theory models (Nigg, Allegrante, & Ory, 2002), seems to indicate that more detailed investigations of motivational constructs (specifically enjoyment) within the field of exercise science are warranted.

The measures of enjoyment compared in this study were chosen based on their use in the physical activity literature, current state of development, and item content. Five measures of enjoyment were selected for evaluation. The scales included the Physical Activity Enjoyment Scale (PACES; Kendzierski & Decarlo, 1991), a modified version of the PACES (PACES-M) developed by Motl et al (2001), the Interest/Enjoyment scale from the Intrinsic Motivation Inventory (IMI-joy; McAuley, Duncan, & Tammen, 1989; Ryan, 1982), the enjoyment scale from the revised Motivation for Physical Activity Measure (MPAMR-joy; Ryan et al., 1997), and the Groningen Enjoyment Questionnaire (GEQ; Stevens, Moget, de Greef, Lemmink, & Rispens, 2000). These scales, which are commonly used to measure enjoyment of exercise and physical activity, were developed from research on sport motivation and performance, intrinsic motivation, and/or flow theory.

The first scale selected was a stand alone scale designed specifically to examine enjoyment of physical activities. The PACES was developed from 39 items that were created based on literature reviews, synonyms and antonyms, personal experience, and interviews with individuals about their attitudes toward physical activity (Kendzierski & Decarlo, 1991). Although no specific definition of enjoyment was presented, item content corresponds to several of the current definitions of enjoyment (Kendzierski et al., 1991; Wankel, 1993; Scanlan & Simons, 1992). The 39 items were rated by three exercise adherence experts, 16 items were selected and 3 additional items were suggested. The 19-item scale was administered to a small number of undergraduates (n = 30) after completing their workout on a stationary exercise bike. The correlation of one item with the scale total score (item-total correlation) was less than 0.30. This item was dropped. The remaining 18 items were then administered to 33 students after an abdominal machine workout. Item-total correlations ranged from 0.45 to 0.87. The final version of the PACES consisted of 18 bipolar statements designed to assess enjoyment of a given physical activity. Kendzierski & Decarlo (1991) reported that changing the instructions to include the activity being studied was acceptable. Similar

74

values of internal consistency reliability have been reported in adults (0.93; Kendzierski & Decarlo, 1991) and children (0.90; Crocker, Bouffard, & Gessaroli, 1995).

The PACES has been used to examine intervention (Calfas et al. 2000) and experimental outcomes (Motl et al. 2000), factors associated with activity class enjoyment (Wininger, 2002; Wininger & Pargman, 2003), and the relationship between social cognitive variables and physical activity (Rovniak et al., 2002). In addition, Kendzierski and Decarlo (1991) reported significantly higher PACES scores during an external focus exercise condition compared to a condition intended to invoke boredom. Despite these findings, the factor structure originally hypothesized (single enjoyment factor) has not been supported in adolescents (Crocker et al., 1995) or young adults (Hales, 2005). After finding the single factor model provided poor fit, Hales (2005) conducted an exploratory analysis of the underlying factor structure of the scale. Results from this analysis and suggestions by Crocker et al. (1995) lead to the development of a two factor model for the PACES. This model included 15 of the original 18 items with 10 items loading on an "enjoyment" factor and 5 items indicating an "energy and achievement" factor. This model was cross-validated on a separate sample of young adults and was found to fit well (CFI = 0.944, RMSEA = .078). Based on this work the 10-item PACES-joy scale was selected for comparison in this study.

Recently the original 18-item PACES was modified for use with adolescent girls involved with a physical activity intervention. The modified version of the PACES (PACES-M) was altered in several ways. The primary changes from the original PACES were to the rating scale and item wording. The 7-point bipolar scale from the PACES was replaced with single statements rated on a 5-point Likert-type scale. The wording was also changed for several items to enhance understanding. Wording changes were based on focus group sessions with adolescent girls, but were not specifically reported (Motl et al., 2001). In addition, two items were dropped. One due to similarities in content with items already within the scale and the other because it was judged irrelevant to enjoyment of

adolescent girls.

The factor validity of the PACES-M has been supported in adolescent girls (Motl et al., 2001), young children (Moore et al., 2004), and young adults (Hales, 2005). The factor structure of the PACES-M includes a single enjoyment factor and a method effect for item structure (positive item uniquenesses allowed to correlate). Support for the factor invariance across gender, race, and over time has also been reported (Hales, 2005; Motl et al. 2001). Although the PACES-M is a direct modification of the PACES no comparisons of the two scales have been reported. Due to the changes made to the rating scale and items, generalization across scales is limited

The third scale selected, the Gronigen Enjoyment Questionnaire (GEQ), was developed as part of an intervention intended to increase leisure-time physical activity in older adults (Stevens, Moget, de Greef, Lemmink, & Rispens, 2000). It was designed to be a short, easy to administer, unidimensional measure of enjoyment. The definition of enjoyment which guided instrument development was based on Csikzentmihalyi's (1975) concept of autotelic experience (one of the nine dimensions of flow). Based on this definition and a search of the literature, 30 items were formed. Ten were selected for the final scale based on an exploratory factor analysis. Factor loadings from the exploratory analysis ranged from 0.58 to 0.86. Coefficient alpha (0.88) and stability reliability of total score (0.84) indicate acceptable internal consistency of items and stability of total score over a two week period.

Evidence from a cross-sectional analysis of 329 sedentary adults (55-65 years old) supported using GEQ scores to discriminate across stage of exercise change (Stevens et al., 2000). In the same study, GEQ scores were found to have small to moderate correlations with perceived ability (0.31-0.44), self-efficacy (0.45-0.55), and social support (0.17-0.29). Approximately 30% of this sample then completed an 18-month active living intervention. GEQ scores were found to discriminate between adherers and drop-outs at 6- and 18-months of the intervention (Stevens et al., 2003).

Development of the Intrinsic Motivation Inventory (IMI) is credited to Ryan (1982), but the instrument has been more thoroughly investigated with respect to exercise and sport by McAuley and colleagues (McAuley, Duncan, & Tammen, 1989; McAuley, Wraith, & Duncan, 1991). The IMI consists of 27-items scored on a seven point Likert-type scale designed to measure five underlying dimensions of intrinsic motivation (interest-enjoyment, perceived competence, effort, pressure-tension, and perceived choice). McAuley et al. (1991) point out that similar items within a subscale, or unneeded subscales, can be removed without changing the internal consistency or factor structure of the remaining items. It is also advantageous that item wording can be changed to fit a variety of activities or situations. The IMI is often used in studies of human behavior to measure intrinsic motivation and self-regulation (McAuley et al., 1989; McAuley et al., 1991; Ryan, Koestner & Deci, 1991). Confirmatory factor analysis results have supported a five subscale version of the IMI worded to measure motivation in aerobic exercise. The scale was administered to beginner, intermediate, and advanced participants in a 10-week aerobic dance class (McAuley et al., 1991). The fit of the five correlated factors model was good (CFI = 0.94, RMSR = 0.057).

The interest/enjoyment subscale (IMI-joy) was selected for comparison in this research study. This scale is considered the self-report measure of intrinsic motivation in the IMI. Acceptable internal consistency reliabilities have been reported for three-, five-, and seven-item versions of the IMI-joy subscale (0.78 – 0.92; Markland & Hardy, 1997; McAuley et al. 1989;1991; Tsigilis & Theodosious 2003). In two studies by McAuley and colleagues (1989; 1991) and a study by Markland and Hardy (1997) the item reflecting "thinking about enjoyment during the activity" was found to be problematic (small factor loading and R^2). Results were often reported with this item excluded.

The final enjoyment scale select for comparison was from the Revised Motivation for Physical Activity Measure (MPAM-R; Ryan et al., 1997). The original version of the MPAM contains 23 items tapping three motives; a body-related factor, a competence factor, and an

interest/enjoyment factor (Frederick & Ryan, 1993). The factor structure was supported in college

students (n = 150) and an adult population (n = 376; mean age 39). The enjoyment/interest subscale

from the original MPAM and MPAM-R are identical except that one item (Because I like the

excitement of participation) has been added to the revised version.

The MPAM-R was developed to differentiate between appearance and fitness goals

contained within the body-related motives scale in the original MPAM and to add a social motives

scale. Revisions were based on findings from two pilot studies. The separation of appearance and

fitness goals and inclusion of social motive is thought to provide additional coverage of intrinsic and

extrinsic factors which may significantly related to participation (Ryan et al., 1997).

The internal consistency reliability of the enjoyment/interest subscale of the MPAM-R

(MPAMR-joy) has been found to be consistently around 0.90 (Wilson, Rodgers, & Fraser, 2002).

Scores have also been reported to have small to moderate correlations with energy expenditure (r =

.30), length of workout (r = 0.23), attendance (r = 0.52), and dropout (r = -0.43). Mean MPAMR-joy

scores have not been found to differ between men and women, but significantly higher scores have

been noted for adherers to an exercise program compared to non-adherers and for individuals

participating in sports compared to fitness exercisers (Frederick & Ryan, 1993; Ryan et al., 1997).

The factor structure of the MPAM-R has been supported by exploratory factor analysis.

Using data from 155 new members at a university fitness center, Ryan et al. (1997) found a five

factor solution similar to the one proposed. The enjoyment factor accounted for more variance than

other subscales with factor loadings ranging from 0.49 to 0.85. Two of the seven enjoyment items

did, however, cross load on the competence factor.

Each of the five measures of enjoyment compared in this study is hypothesized to represent

a single factor. They have been used as indicators of enjoyment of physical activity, sport, and

exercise in a variety of situations, but have never been compared. Because there is currently no

standardized instrumentation for research related to enjoyment of physical activity, or exercise behavior it is important that scales used to indicate enjoyment are comparable and have sound psychometric properties. Therefore the purpose of this research was to evaluate several measures of enjoyment with respect to their comparability, factor validity, internal consistency reliability, and construct validity. To address these issues we: (1) examined the psychometric and distributional properties of each scale, (2) tested the factor structure of each scale, (3) assessed convergence and presence of a method effect, and (4) estimated and compared the relationship of each enjoyment scale with measures of physical activity, perceived competence, and extrinsic participation motives. Based on evidence from the physical activity, exercise, sport, motivation, and enjoyment literature, several hypotheses were developed with respect to the construct validity evidence for the enjoyment scores. First, it was hypothesized that the enjoyment scales would be highly correlated and that a confirmatory model including all scales would support the convergence of items. It was also expected that enjoyment would significantly correlate (0.20 – 0.40) with total physical activity, but not work-related physical activity. Finally, based on cognitive evaluation theory, it was predicted that enjoyment would exhibit a stronger relationship with an intrinsic motive (i.e. perceived competence) than with more extrinsic motives (i.e. social, appearance, and fitness).

Method

<u>Participants</u>

Participants (N = 1023) were recruited from basic physical education classes and the psychology research pool at two universities (n = 733; n = 290). They were informed of the study through direct class contact or a listing of available experiments. Psychology research participants received class credit for involvement in the study. The sample had a mean age of 19.60 ± 1.55 years with a racial distribution of 73.9% white, 18.8% black, and 7.3% other. Most participants were

freshman (n=405), but sophomores (n = 238), juniors (n =204) and seniors (n=174) also completed

the survey. Sixty-eight percent of the participants were female and 27.6% were recruited from

physical education classes.

Measures

PACES-enjoy. The PACES-enjoy consists of 10-items from the original PACES. The items

were found to load on a single factor in an exploratory analysis and cross-validated in a separate

sample using confirmatory factor analysis (Hales, 2005). Each item consists of two bipolar

statements with a seven point scale defining the continuum between them. Respondents were asked

to rate how they felt at the moment about the physical activities in which they usually participate.

PACES-M. The PACES-M consists of 16-statements rated on a 5-point Likert-type scale

ranging from (1) "Disagree a lot" to (5) "Agree a lot". Each statement beings with the phrase "When

I am active,…". The scale was based on the original version of the PACES.

MPAM-R. The MPAM-R consists of 30 items rated on a 7-point Likert-type scale ranging

from (1) "not at all true for me" to (7) "very true for me". It contains five subscales measuring

fitness, appearance, competence, enjoyment, and social motives for physical activity.

Intrinsic Motivation Inventory (IMI). Responses to the interest/enjoyment (IMI-joy) and

perceived competence (IMI-pc) subscales of the IMI were obtained. The IMI-joy subscale contains

seven items rated on a 7-point scale defined by (1) "not at all true", (4) "somewhat true", and (7)

"very true". The perceived competence scale includes six items. Item content was modified to reflect

general physical activities (i.e. "I enjoy participating in physical activities very much"). Internal

consistency reliabilities ranging from 0.78 to 0.92 have been reported for the subscales (Markland &

Hardy, 1997; McAuley et al. 1989;1991; Tsigilis & Theodosious 2003). In addition, the stability of

IMI subscale scores has been assessed over one week using a Greek version of the scale. Single

measure intraclass correlation coefficients of 0.86 and 0.61 for the IMI-joy and IMI-pc were reported. (Tsigilis & Theodosiou, 2003).

Groningen Enjoyment Questionnaire. The GEQ includes 10 items (e.g. "Doing leisure-time physical activity makes me feel good") rated on a 5-point Likert-type scale ranging from (1) "I agree" to (5) " I Disagree" (Stevens et al., 2000). It was designed to be a short, easy to administer, unidimensional measure of enjoyment. Internal consistency of items (0.88) and stability (R = 0.84) of total score over two weeks indicate acceptable scale reliability.

Physical activity. Three measures of physical activity were collected. The International Physical Activity Questionnaire-short form (IPAQ) is a 4-item self-administered survey (Craig et al., 2003). It was developed as part of an international effort to assess population levels of physical activity across countries. In the IPAQ, respondents are asked to quantify the days per week, and hours/minutes per day of vigorous, moderate, walking, and sitting activity. In the current investigation, estimates of total, vigorous, and moderate physical activity were calculated and expressed as MET-min/week. Standardized procedures were followed for data cleaning, analysis, and calculation of scores (Craig et al., 2003; also see www.ipaq.ki.se). The stability reliability of IPAQ scores over one-week calculated for samples from the United States (n = 77) and United Kingdom (n=151) ranged from 0.66 to 0.88. The pooled reliability for the IPAQ across forms (long and short) and countries was 0.76 (95% CI 0.73-0.77; Craig et al. 2003). Concurrent validity estimates (long vs. short form) ranged from 0.61 to 0.89 (median = 0.72) for the United States sample.

Responses to the four items from the Lipid Research Clinic Physical Activity Questionnaire were also obtained (Ainsworth, Jacobs, & Leon, 1993). This measure consisted of two yes-or-no and two likert-type items. A standard scoring protocol was used to classified participants as highly active, moderately active, low active, or very low active. Estimates of stability reliability for this measure

have consistently been above 0.80. Scores from this measure have also been shown to moderately correlated with other measures of physical activity, VO2 max, and treadmill time (Ainsworth, Jacobs, & Leon, 1993; Albanes, Conway, Taylor, Moe, & Judd, 1990; Jacobs, Ainsworth, Hartman, & Leon, 1993).

Occupational activity was assessed using the work index from the Baecke Questionnaire of Habitual Physical Activity (Baecke, Burema, & Frijters, 1982). This index includes eight items related to work activities (i.e. standing, lifting) and strenuousness of work (i.e. sweat, tired). Items were averaged to form a work index. Stability reliability estimates for the work index are generally above 0.80.

Procedures

Data were collected by web-based survey. The web-survey was constructed with help from the university Survey Research Center. Each survey contained an informed consent section, demographic questions (age, sex, etc.), measures of enjoyment and physical activity, and a social desirability scale. Each questionnaire (i.e. PACES-M, IPAQ, etc.) was displayed as a separate section within the web-survey and replicated to match the layout and response format of the original paper-and-pencil version. If needed, instructions were modified to reflect enjoyment of general physical activity (including any physical activity, play, sport, or exercise). The average time it took to complete the survey was 16.35 ± 9.3 minutes. A comparison of responses to web-based and paper-and-pencil versions of the survey was conducted prior to data collection. No meaningful differences were found between the two methods of administration.

Once enrolled in the study, students were contacted via email. Each email message contained a brief overview of the study, a hyperlink to the survey web-site, and a passcode. Passcodes were used to limit access to recruited participants. To enhance response rate, follow-up email messages

82

were sent one and two weeks after the initial contact to remind people to complete the survey. All procedures were approved by the institutional review boards at the universities were data collection took place.

Data analysis

Means, standard deviations, measures of distribution, Pearson product moment correlations, and internal consistency reliability coefficients were calculated using SPSS 13.0. A one-way ANOVA with a Bonferroni correction was utilized to compare enjoyment scores across levels of physical activity. Alpha was set at 0.05. PRELIS (Joreskog & Sorbom, 1996b) was utilized to calculate tests of multivariate normality. Initial tests of the hypothesized factor structure for each enjoyment scale were conducted using confirmatory factor analysis. All confirmatory factor analysis models were tested using AMOS 5.0 with full-information maximum likelihood estimation (Arbuckle, 2003; Arbuckle & Wothke, 1999). Only 2.9% of the 51,150 (50 items x 1023 participants) item responses were missing. Fit indices, parameter estimates, and modification indices for each of the confirmatory factor analysis models were also obtained using LISREL 8.5 (Joreskog & Sorbom, 1996a). When the initial model for a measure of enjoyment did not provide adequate fit to the data, modification indices, parameter estimates, squared multiple correlations, residuals, items content, and previous results were used to identify the most likely areas of misfit.

The root mean square error of approximation (RMSEA), non-normed fit index (NNFI), comparative fit index (CFI), expected cross-validation index (ECVI), and χ^2 statistic were used to evaluate and compare model fit. The χ^2 statistic is very sensitive to sample size. Because large samples are required to obtain stable parameter estimates in confirmatory factor analysis, this statistic indicates rejection of the hypothesized model when discrepancies between the sample and reproduced covariance matrix are quite small (Bollen, 1989; Joreskog, 1993). For this reason, it is

reported but supplemented by other fit indices to draw specific conclusions about model fit. The RMSEA is a standardized estimation that represents closeness of fit of the population data to the model. It is widely used and is considered one of the most informative fit criteria (Byrne, 1998). Values around 0.06 reflect close fit, while values up to and around 0.10 indicate reasonable fit of the model (Browne & Cudeck, 1993; Hu & Bentler, 1999). The 90% confidence interval around the RMSEA point estimate should contain 0.06 or zero to indicate the possibility of close or exact fit. For all multiple groups analyses the RMSEA estimate and CI were multiplied by $\sqrt{2}$ as suggested by Steiger (1998). The CFI and NNFI test the proportionate improvement in fit by comparing the target model to some baseline model (Bentler & Bonett, 1980; Hu & Bentler, 1999). The NNFI, but not the CFI, is affected by model parsimony (more complex models are penalized). Values for the CFI and NNFI approximating 0.95 indicate good fit (Hu & Bentler, 1999). Acceptable model fit was based on CFI and NNFI values of 0.90 (Bentler & Bonett, 1980; Hu & Bentler, 1999). The ECVI takes into account both model fit and parsimony and reflect how well a model might cross-validate in similar samples. A lower value for the ECVI indicates a model with fit and parameter values most likely to replicate in future samples (Browne & Cudeck, 1989; Hu & Bentler, 1998, Bandalos, 1993). Overall judgment of model fit was based on model interpretability, the cumulative evidence from all measures of fit, and current suggestions that strict "cut-off" values for specific indices may not always be appropriate (Byrne, 1998; Hu & Bentler, 1998; Marsh, Hau, & Wen, 2004).

The measurement invariance between men and women (gender) and black and white (race) participants was also examined in the best fitting IMI-joy, MPAM-joy, and GEQ models. To accomplish this, two models were tested and compared between men and women and then between black and white participants. In the first model, the factor structure was specified to be the same in both groups, but parameter values were not forced to be equal (configural invariance). In the other model, the factor loadings were constrained to be equal across groups (metric invariance). The

equivalence of these two models is often considered sufficient evidence for concluding invariance across group. If the RMSEA, NNFI, and CFI were the same, or very similar, in these two models it was concluded that constraining the factor loadings to be equal between groups did not affect model fit. Examining differences in these three fit indices has been found to work well for judging measurement invariance (Cheung & Rensvold, 2002). The invariance of the PACES and PACES-M has been previously established in this sample (Hales, 2005).

To examine the convergent validity of the enjoyment scales and possible method effect for items within a scale, a series of confirmatory factor analysis models were tested and compared (see figure 4.1). In confirmatory factor analysis two models are considered nested if a simpler model can be obtained by imposing a set of restrictions onto a more complex model. The only difference between the models is the number of free and fixed parameters estimated (i.e. the nested, or simpler, model has more fixed, or constrained parameters). The first model (M1) included a single enjoyment factor indicated by each of the 50 items. The second model (M2) included five correlated factors representing enjoyment measured by each scale. In this model, each factor was indicated by the items from its respective scale. The next model (M3) was a combination of M1 and M2, including a single enjoyment factor indicated by each of the 50 items and a method factor for each scale indicated by the items from that scale only. Comparison of M3 and M2 was used as the primary test of the method effect. Due to the similarity in item wording across scales, a final model (M4) including correlated uniquenesses among similarly worded items was tested. The χ^2 difference tests were calculated and reported, but because they are sensitive to sample size, comparisons of model fit were based primarily on changes in the values of the RMSEA, NNFI, and CFI between nested models. Examining differences in these three measures has been found to be superior to interpretations based strictly on χ^2 difference tests (Cheung & Rensvold, 2002). An absolute change in CFI of .01 between nested models has been reported to work well for testing differences (Cheung

& Rensvold, 2002). Subjective assessments of RMSEA point estimates and 90% confidence intervals between nested models were also used to judge meaningful change in fit between models. Based on current recommendations parameter estimates, standard errors, z-values, and squared multiple correlations were inspected for sign and magnitude (Boomsma, 2000; Joreskog, 1993).

Results

Descriptive statistics and measures of distribution for each of the enjoyment scales are presented in Table 4.1. For each enjoyment scale, the mean score is above the midpoint of the scale and all distributions are negatively skewed indicating positive endorsement for all scales. The skewness and kurtosis values are all within an acceptable range with only the kurtosis for the GEQ exceeding 1.0. Skewness and kurtosis values were also obtained for the individual items included in each scale. The absolute value of kurtosis and skewness for the items ranged from 0.008 to 5.199 and 0.094 to 2.22, respectively. Of the 50 items, eight had kurtosis values greater than two (one item > 5; one item 3-4; six items 2-3). The mean absolute value of the kurtosis was 0.918 with a relative multivariate kurtosis value of 1.273. One item had a skewness value greater than 2.0. This item, "It makes me depressed", from the PACES-M, also had the largest kurtosis value.

Also presented in Table 4.1 are estimates of internal consistency reliability and the item-total correlations for the scales. Item-total correlations represent the relationship between an item and the total score for a scale. The item-total correlations ranged from 0.464 to 0.891 with most values above 0.70. The internal consistency reliabilities exceed 0.90 for all scales.

Space constraints limit the presentation of all items from each scale, but a general overview and subjective comparison of scale content is necessary. The number of items included in the scales varied from 7 to 16. Four of the instruments have items with single statements and Likert-type rating scales with five or seven response options. The remaining measure (PACES-joy) is bipolar,

with a seven point scale defining the continuum between statements. A simple content analysis was conducted grouping items with similar wording and descriptors across measures of enjoyment. For example, the items "When I am active, I feel bored." and "I think physical activities are boring" are very similar in content and were grouped together. The descriptors "enjoy", "like/dislike", and "fun" were used in four of the five measures and "interest/bored" was used in every measure. In fact, of the 50 items 22 had one of those descriptors. Of the 50 items, nine appear in only one measure. Four of these items are part of the GEQ. These four items include concepts related to effort, loss of time, relaxation, and being ones self and do not seem to reflect the common conceptualization of enjoyment of physical activities.

CFA for Individual Enjoyment Measures

The factor validity of the five measures of enjoyment was examined by conducting separate confirmatory factor analyses for each instrument. A single factor solution was hypothesized for all scales. In each confirmatory model, the metric was set by fixing the factor loading of one item to 1.0. Based on previous research the hypothesized model for the PACES-M included correlated uniquenesses among the positively worded items (Hales, 2005; Moore et al., 2004; Motl et al. 2001). The fit indices for the confirmatory factor analyses testing the measurement model of each scale are presented in Table 4.2.

The NNFI and CFI indicate good fit for the PACES-M, PACES-enjoy, and MPAMR-joy models. The RMSEA was less than 0.08 for the PACES-M and PACES-enjoy, but around 0.10 for the MPAMR-joy. For the GEQ and the IMI-joy scales the initial single factor model did not fit the data well. For both scales the NNFI and CFI were lower than expected and the RMSEA was greater than 0.12. Modification indices were examined to determine the most likely areas of misfit. For the IMI-joy scale a very large modification index (> 300) was found for the error covariance between

the two negatively worded items in the scale. When this single correlated uniqueness was included, fit of the model improved substantially. This modification was retained in the final model of the IMI-joy scale. Examination of the GEQ model revealed misfit for items 7 to 10. The large modification indices for the six error covariances among these items indicated substantial common variance not accounted for by the enjoyment factor. Based on this finding and the content analysis, a second model was specified which allowed all error covariances among items 7 to 10 to be freely estimated. The CFI for this model indicated good fit, while the NNFI and RMSEA suggested adequate fit to the data.

The fit indices used to examine the invariance of factor structure and factor loadings are also presented in Table 4.2. There was support for invariance between male and female and between black and white participants. The fit of the models setting the factor structure equivalence between races (RI 1) and between genders (GI 1) was acceptable for all scales. When factor loadings were constrained between races (RI 2) and between genders (GI 2) values of the NNFI, CFI, and RMSEA changed very little. The largest decrease in CFI was 0.002 and increases in χ^2 were not significant for any of the scales.

Convergent Validity and Tests of Method effects

The Pearson product moment correlation coefficients among enjoyment scales are presented in Table 4.3. All correlations were significantly greater than zero, ranging from 0.539 to 0.825. The mean correlation among enjoyment scales was 0.689 (95% CI = 0.653 – 0.722). To examine the convergent validity of the enjoyment scales and possible method effect for items within a scale a series of confirmatory factor analysis models were tested and compared (see Figure 2). Fit indices for this series of models and model comparisons can be found in Table 4.4. The single enjoyment factor, M1, provided poor fit to the data (CFI = 0.740, NNFI = 0.701, RMSEA = 0.093). The fit of

this model indicates that a single factor does not adequately account for the covariance among the enjoyment items. Fit for M2, five correlated factors, was acceptable. Correlations among the factors ranged from 0.550 to 0.874. The model hypothesized to fit best, M3, included both the general enjoyment factor and method factors for each scale. The CFI and NNFI indicated adequate fit for M3, while the RMSEA suggests good fit of the model to the data. All of the loadings for the enjoyment factor were significant, ranging from 0.394 to 0.838. Loadings for the five method factors were all negative. Four loadings were not significantly different than zero. Models M2 and M3 fit substantially better than M1, and M3 fit better than M2 (ΔCFI = 0.014). The addition of autocorrelated errors (correlation among like items across scales) in M4 resulted in a good fitting model (CFI = 0.946, NNFI = 0.934, RMSEA = 0.044). Of the 53 correlated errors included in this model 26 were not statistically different from zero. Change in CFI, NNFI, and RMSEA indicated that model fit improved substantially from M1 to M4.

Further evidence for the construct validity of the enjoyment scales was provided by correlating each enjoyment scale with estimates of vigorous, moderate, total and work physical activity and comparing mean enjoyment scores across four levels of physical activity. Correlations of the total score for each enjoyment scale with the physical activity estimates can be found in Table 4.5. As predicted the correlations between work activity and enjoyment were very small ($\leq$ 0.10), while correlations with vigorous activity were larger (0.271 – 0.340). Correlations with vigorous activity were similar for all enjoyment scales, ranging from 0.27 to 0.34. Correlations with total physical activity were of similar magnitude across scales (0.223 – 0.286) and slightly lower than those for vigorous activity. Correlations with moderate activity were small, but generally larger than those for work activity.

To compare mean enjoyment score across levels of physical activity, four groups were formed from responses to the Lipid Research Clinic Physical Activity Questionnaire. Participants

were classified as very low active, low active, moderately active, or highly active. A one-way ANOVA with a Bonferroni correction for multiple comparisons was conducted for each enjoyment scale. Mean scores for each group are presented in Table 4.6. A similar pattern of results emerged for each enjoyment scale. Scores for each scale increased as activity level increased. The mean difference between very low and low active groups was not significant for any enjoyment scale, while the highly active group had significantly higher enjoyment scores compared to each of the other activity groups. Mean enjoyment scores for the moderately active group were significantly higher than scores for the very low active group, but mean differences with the low active group varied across enjoyment scale. A plot of mean enjoyment scores across levels of physical activity is presented in Figure 4.2. Due to the similarity in results from the ANOVA analysis, post-hoc regression analyses were conducted to determine if the rate of change in enjoyment score as physical activity increased was similar across enjoyment scales. The slope of the regression lines ranged from 0.308 to 0.385. This difference was not statistically significant. The adjusted R^2 values ranged from 0.094 to 0.148.

Construct validity evidence for enjoyment scale score interpretations was also examined by calculating the correlation between each measure of enjoyment and each measure of perceived competence and between each measure of enjoyment and the measures of extrinsic motives for physical activity (fitness, social, and appearance). These correlations are presented in Table 4.5. As hypothesized, correlations with the intrinsic motive (perceived competence) were moderate and consistently larger than correlations with extrinsic motives. Correlations between enjoyment and a given motivation were also very similar in magnitude across enjoyment scales. Most correlations with perceived competence were between 0.537 and 0.675. Two exceptions should be noted. The correlations between MPAM-joy and MPAM-pc (0.847) and between IMI-joy and IMI-pc (0.775)

were substantially, and significantly, higher. This seems to indicate a possible method effect that is also evident in the correlations of MPAM-joy with MPAM-fit (0.523) and MPAM-social (0.520).

Finally, the relationship between total score for each enjoyment scale and a measure of social desirability was calculated. Correlations ranged from 0.045 to 0.151. These correlations indicate a weak relationship between measures of enjoyment and the tendency for social desirable response sets.

Discussion

The purpose of this research was to evaluate several measures of enjoyment with respect to their comparability, factor validity, internal consistency reliability, and construct validity. In general, the results of the confirmatory factor and correlational analyses support the similarity of enjoyment scale scores. The invariance analysis provides additional evidence for the comparability of enjoyment scores between black and white and male and female young adults. The results indicate that the factor structure and loadings were equivalent between groups for all enjoyment scales. The hypothesized relationships of physical activity enjoyment with physical activity, perceived competence, and extrinsic motives were also supported. The magnitude of these relationships was very similar across the five measures of enjoyment. Results from the confirmatory factor analyses do, however, suggest that there is significant common variation among items on a given scale even after accounting for the enjoyment factor.

A simple one-factor model provided adequate to good fit for two of the five measures of enjoyment. For the PACES-M, GEQ, and IMI-joy a single enjoyment factor did not adequately explain the covariance among the items. The models for these scales required one or more correlated uniquenesses to improve fit. In two cases, PACES-M and IMI-joy, the correlated uniquenesses were interpreted to represent a method effect. For the PACES-M, a single-factor

model including correlated uniquenesses among positively worded items was hypothesized based previous research (Hales, 2005; Moore et al., 2004; Motl et al., 2001). The factor validity and invariance of this model have been supported in adolescent girls, young adults, and third grade students. The significant correlated uniqueness between the two negatively phrased items on the IMI-joy scale was not anticipated, but was judged to be a function of item structure (negative items) and item content ("activities are boring" and "activities do not hold my attention").

The correlated uniquenesses included in the best fitting model of the GEQ could not be attributed to a method effect. In this model, the correlations among the uniquenesses for items 7 to 10 were freely estimated. These items relate to effort, loss of time, relaxation, and being ones self. Based on the content analysis and theoretical basis for the GEQ, it is hypothesized that these items may reflect some aspect of an autotelic experience that is related, but distinct from enjoyment. The only previous support for the factor structure of the GEQ is from an exploratory factor analysis conducted on the 30 items constructed during scale development. In this analysis, the 10 items included in the final version of the GEQ loaded on a single factor. The rotated factor loadings were lowest for items 7 to 10. No confirmatory or exploratory factor analyses using the 10 items selected for the final scale have been reported. The four items in question do seem to reflect Csikszentmihalyi's (1975) definition of autotelic experience, which was used to guide item development. There is some debate about whether the terms enjoyment, autotelic experience, and flow are synonymous (Kimiecik & Harris, 1996; Wankel, 1997). It is evident that the factor structure of the scale needs to be examined further. Interestingly, the bivariate correlations of the GEQ with the other enjoyment measures were not systematically lower than those among the PACES-M, PACES-enjoy, IMI-joy, and MPAM-joy.

The fit indices for the one-factor models of the MPAM-joy and PACES-enjoy are generally good. The RMSEA for the MPAM-joy suggests errors of approximation in the population that are

larger than desired (Browne & Cudeck, 1993; MacCallum, Browne, & Sugawara, 1996). The CFI, NNFI, ECVI, parameter estimates, and squared multiple correlations do, however, support the factor validity of the scale. Based on the confirmatory factor analysis results, the PACES-enjoy scale is judged to have the strongest evidence for factorial validity. This is based on the fact that the single–factor model provides a good representation of the covariance among the items and does not include correlated uniquenesses, making it easier to interpret. The PACES-enjoy measure was developed and cross-validated with random samples drawn from the data used in this analysis. Despite, strong support for the factorial validity and invariance of the scale (Hales, 2005), confirmation of these findings is needed.

The results from the confirmatory factor analyses testing the convergence of enjoyment measures indicate that a single factor did not adequately account for the covariance among items from all enjoyment scales. When each item was allowed to load on a method factor for that scale and the enjoyment factor, fit improved substantially. Method effects are thought to exist when the "true" relationship among items or variables is obscured by some other similarity among the items or variables (i.e. items measured within same instrument). Additional insight into the effect of method is inferred from the results of the correlational analysis. In most cases, the method effect seems to be an artifact that does not change the relationship of enjoyment with other measures. However, the relationship between enjoyment and perceived competence is substantially higher when the two constructs are measured with subscales of the same instrument (i.e. IMI or MPAM-R). In this sample, measuring the two intrinsic motives with the same instrument inflates the relationship.

An alternate explanation for the method effect could be item proximity rather than the effect of a certain instrument. What this implies is that a response pattern could have emerged for items read one after the other, or items on the same page. No direct test comparing these alternatives is

possible using these data, but the latter was most likely minimized during data collection because items within a scale were grouped (5 to 10 items per group) and randomly administered. This randomization resulted in all scales being completed in two or three nonconsecutive parts. This could have resulted in a response set for a group of items, but would not account for a response set across all items within a scale. Both types of method effect can obscure the correlation between variables and should be considered when examining the relationships among motivations for physical activity.

The results of the correlation and ANOVA analyses provide additional construct validity evidence for the measures of enjoyment. The relationships between enjoyment and physical activity reported in the literature are generally around 0.22 (Hales, 2005). This is consistent with the findings for total and vigorous physical activity and for the mean differences in enjoyment scores across physical activity levels. The relationships to moderate physical activity are small, indicating that in this sample moderate activity is minimally related to enjoyment. This was somewhat unexpected, but may reflect activity choices made by college students. Most moderate-intensity activities that college students report may be activities of daily living (i.e. cleaning house, getting to class, walking to the bar). These types of activities should, theoretically, be unrelated to enjoyment of physical activities such as play, sport, or exercise. Similarly, work related activity had small relationships with enjoyment. This was expected. Work related activity is often required, which, from the perspective of self-determination theory, should be less enjoyable due to lack of control over participation (Deci & Ryan, 1985). Finally, the relationships of enjoyment with intrinsic and extrinsic motives were as hypothesized. These findings support the proposition that enjoyment should be more highly related to perceived competence (another intrinsic motive) compared to extrinsic motives for participation in physical activities (Deci and Ryan, 1985; Ryan, 1982).

The relationships among enjoyment scales are moderate to large, generally supporting the convergent validity of the measures. Factor correlations range from 0.550 to 0.874 and the mean bivariate correlation among scales is 0.689. There are no universally accepted criteria for judging convergence, but similar correlations have been reported in studies comparing multiple measures of physical self concept (Marsh, Richards, Johnson, Lawrence, & Tremayne, 1994) and goal orientations (Jagacinski & Duda, 2001). In addition, the relationships between enjoyment and physical activity, perceived competence, and extrinsic motives are very similar across enjoyment scales. Overall, these finding support the construct validity and comparability of scores from the PACES-M, PACES-enjoy, IMI-joy, MPAM-joy, and GEQ.

Although it would be nice to definitively suggest a "best" measure of enjoyment, the results indicate that the enjoyment scales studied are somewhat interchangeable. Based on these findings the following observations are provided:

(1) The seven-item scales functioned just as well as the 16-item scale. Considering the time constraints of data collection, this may be important to consider.

(2) The content of most enjoyment scales is similar. The descriptors "enjoy", "like/dislike", "fun", and "interest/bored" are used most often to create items. These terms reflect the most common definitions and descriptions of enjoyment in the exercise, sport, and physical activity literature (Scanlan & Simons, 1992; Wankel, 1993; Wankel & Kreisel, 1985).

(3) For each questionnaire, more than eighty percent of participants have enjoyment scores above the scale midpoint. While only 40 to 60% of them would be considered regularly active. This indicates that most participants enjoy physical activities, or think they are supposed to enjoy physical activity, even if they are not active. The general positive endorsement of each scale along with the fact that mean enjoyment scores did not differ between very low and low active individuals suggests

a floor effect for these measures of enjoyment. Items that discriminate better across levels of physical activity may need to be developed.

(4) Method effects should be considered when selecting a measure of enjoyment and when ordering items within a research survey. If motivations to be physically active are measured by subscales of the same instrument, method effects may obscure the "true" relationship among the constructs.

(5) Researchers should be aware of the construct definition that was used during scale development. When reporting results, it should be made clear what construct an instrument was designed to measure and the definition of that construct.

The purpose of this research was to compare and extend the construct validity evidence for several measures of physical activity enjoyment. Scores from the five measures of enjoyment are comparable. Slight differences in factor structure and the presence of method effects do not affect the relationship of any enjoyment scale with physical activity. The method effect does inflate the relationship between constructs measured with subscales from the same instrument (i.e IMI and MPAM-R). The measurement properties of the scales are similar for males and females and for black and white young adults. Future research should extend these findings to determine if the scales are invariant across a wide range of ages. If they are invariant, scores could be used to compare enjoyment of physical activity across the lifespan and to compare changes in enjoyment over longer periods of time. Another interesting finding that requires further study is that the length of the instrument did not affect the results. This suggests a need to determine the most economical measure that fully represents the enjoyment construct. If a three- or four-item scale is found to be equivalent to a 10- or 18-item scale, use of the short scale would be preferred in a number of situations. Overall, this research supports the construct validity and similarity of scores from the PACES, PACES-M, GEQ, IMI-joy, and MPAM-R-joy.

References

Ainsworth, B. E., Jacobs, D. R., & Leon, A. S. (1993). Validity and reliability of self-reported physical-activity status - The lipid-research-clinics questionnaire. <u>Medicine and Science in Sports and Exercise, 25</u>, 92-98.

Ajzen, I. (1988). <u>Attitudes, personality and behavior.</u> Chicago: Dorsey.

Ajzen, I., & Fishbein, M. (1974). Factors influencing intentions and the intention-behavior relation. <u>Human Relations, 27</u>(1), 1-15.

Albanes, D., Conway, J. M., Taylor, P. R., Moe, P. W., & Judd, J. (1990). Validation and comparison of eight physical activity questionnaires. <u>Epidemiology, 1</u>, 65-71.

American Educational Research Association (1999). <u>Standards for educational and psychological testing.</u> Washington, DC: AERA.

Arbuckle, J. L. (2003). <u>Amos 5.0 update to the Amos user's guide</u>. Chicago, IL: SPSS

Arbuckle, J. L., & Wothke, W. (1999). <u>Amos 4.0 user's guide.</u> Chicago, Il: SmallWaters Corporation.

Baecke, J. A. H., Burema, J., & Frijters, J. E. R. (1982). A short questionnaire for the measurement of habitual physical activity in epidemiological studies. <u>American Journal of Clinical Nutrition, 36</u>, 936-942.

Bandalos, D. L. (1993). Factors influencing cross-validation of confirmatory factor-analysis models. <u>Multivariate Behavioral Research, 28</u>(3), 351-374.

Bandura, A. (1986). <u>Social foundations of thought and action</u>. Englewood Cliffs, NJ: Prentice-Hall.

Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance-structures. <u>Psychological Bulletin, 88</u>(3), 588-606.

Bollen, K. A. (1989). <u>Structural equations with latent variables.</u> New York: John Wiley.

Boomsma, A. (2000). Reporting analyses of covariance structures. <u>Structural Equation Modeling, 7,</u> 461-483.

Browne, M. W., & Cudeck, R. (1989). Single sample cross-validation indexes for covariance-structures. Multivariate Behavioral Research, 24(4), 445-455.

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), Testing structural equation models (pp. 136-162). Newbury Park, CA: Sage.

Byrne, B. M. (1998). Structural equation modeling with LISREL, PRELIS, and SIMPLIS: Basic concepts, applications, and programming. Mahwah, N.J.: Lawrence Erlbaum.

Calfas, K. J., Sallis, J. F., Nichols, J. F., Sarkin, J. A., Johnson, M. F., & Caparosa, S. (2000). Project GRAD: Two-year outcomes of a randomized controlled physical activity intervention among young adults. American Journal of Preventive Medicine, 18(1), 28-37.

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. Structural Equation Modeling, 9(2), 233-255.

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

Crocker, P. R. E., Bouffard, M., & Gessaroli, M. E. (1995). Measuring enjoyment in youth sport settings - A confirmatory factor-analysis of the physical-activity enjoyment scale. Journal of Sport & Exercise Psychology, 17(2), 200-205.

Csikszentmihalyi, M. (1975). Beyond boredom and anxiety: Experience flow in work and play. San Francisco: Jossey-Bass.

Deci, E. L., & Ryan, R. M. (1985). Intrinsic motivation and self-determination in human behavior. New York: Plenum.

DiLorenzo, T. M., Stucky-Ropp, R. C., Vander Wal, J. S., & Gotham, H. J. (1998). Determinants of exercise among children. II. A longitudinal analysis. Preventive Medicine, 27(3), 470-477.

Dishman, R. K., Motl, R. W., Saunders, R., Felton, G., Ward, D. S., & Dowda, M. (2005). Enjoyment mediates effects of a school-based physical-activity intervention. <u>Medicine and Science in Sports and Exercise, 37</u>(3), 478-487.

Dishman, R. K., & Sallis, J. F. (1994). Determinants and interventions for physical activity and exercise. In C. Bouchard, R. J. Shephard & T. Stephens (Eds.), <u>Physical activity, fitness, and health: International proceedings and consensus statement</u> (pp. 214 - 238). Champaign, IL: Human Kinetics.

Frederick, C. M., & Ryan, R. M. (1993). Differences in motivation for sport and exercise and their relations with participation and mental health. <u>Journal of Sports Behavior, 16</u>, 124-146.

Garcia, A. W., & King, A. C. (1991). Predicting long-term adherence to aerobic exercise: A comparison of two models. <u>Journal of Sport and Exercise Psychology, 13</u>, 394 - 410.

Hales, D. P. (2005). <u>Factor validity, invariance, comparison of several measures of physical activity enjoyment</u>. Unpublished Dissertation, University of Georgia, Athens.

Hu, L. T., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. <u>Psychological Methods, 3</u>(4), 424-453.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. <u>Structural Equation Modeling, 6</u>(1).

Izard, C. E. (1977). <u>Human emotions.</u> New York: Plenum.

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

Jagacinski, C. M., & Duda, J. L. (2001). A comparative analysis of contemporary achievement goal orientation measures. <u>Educational and Psychological Measurement, 61</u>(6), 1013-1039.

Joreskog, K. G. (1993). Testing structural equation models. In K. A. Bollen & J. S. Long (Eds.), Testing structural equation models. Newbury Park, CA: Sage.

Jöreskog, K. G., & Sörbom, D. (1996a). LISREL 8: User's reference guide. Chicago, IL: Scientific Software International.

Joreskog, K., & Sorbom, D. (1996b). PRELIS 2: Users reference guide. Hillsdale, NJ: Lawrence Erlbaum.

Kendzierski, D., & Decarlo, K. J. (1991). Physical activity enjoyment scale - two validation studies. Journal of Sport & Exercise Psychology, 13(1), 50-64.

Kimiecik, J. C., & Harris, A. T. (1996). What is enjoyment? A conceptual/definitional analysis with implications for sport and exercise psychology. Journal of Sport & Exercise Psychology, 18(3), 247-263.

MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. Psychological Methods, 1, 130-149.

Markland, D., & Hardy, L. (1997). On the factorial and construct validity of the intrinsic motivation inventory: Conceptual and operational concerns. Research Quarterly for Exercise and Sport, 68(1), 20-32.

Marsh, H. W., Hau, K., & Wen, Z. (2004). In search of golden rules: comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. Structural Equation Modeling, 11(3), 320-341.

Marsh, H. W., Richards, G. E., Johnson, S., Lawrence, R., & Tremayne, P. (1994). Physical self-description questionnaire: Psychometric properties and a multitrait-multimethod analysis of relations to existing instruments. Journal of Sport and Exercise Psychology, 16, 270-305.

Mcauley, E., Duncan, T., & Tammen, V. V. (1989). Psychometric properties of the intrinsic motivation inventory in a competitive sport setting - A confirmatory factor-analysis. Research Quarterly for Exercise and Sport, 60, 48-58.

McAuley, E., Wraith, S., & Duncan, T. E. (1991). Self-efficacy, perceptions of success, and intrinsic motivation for exercise. Journal of Applied Social Psychology, 21, 139-155.

Messick, S. (1995). Validity of psychological-assessment - Validation of inferences from persons responses and performances as scientific inquiry into score meaning. American Psychologist, 50(9), 741-749.

Moore, J. B., Yin, Z., & Gutin, B. (2004). Enjoyment of physical activity in children: Instrument validation in the MCG FitKid project. Journal of Sport & Exercise Psychology, 26, S138-S139.

Motl, R. W., Berger, B. G., & Leuschen, P. S. (2000). The role of enjoyment in the exercise-mood relationship. International Journal of Sport Psychology, 31(3), 347-363.

Motl, R. W., Dishman, R. K., Saunders, R., Dowda, M., Felton, G., & Pate, R. R. (2001). Measuring enjoyment of physical activity in adolescent girls. American Journal of Preventive Medicine, 21(2), 110-117.

Nigg, C. R., Allegrante, J. P., & Ory, M. (2002). Theory-comparison and multiple-behavior research: common themes advancing health behavior research. Health Education Research, 17(5), 670-679.

Rovniak, L. S., Anderson, E. S., Winett, R. A., & Stephens, R. S. (2002). Social cognitive determinants of physical activity in young adults: A prospective structural equation analysis. Annals of Behavioral Medicine, 24(2), 149-156.

Ryan, R. M. (1982). Control and information in the intrapersonal sphere: An extension of cognitive evaluation theory. Journal of Personality and Social Psychology, 43, 450-461.

Ryan, R. M., Frederick, C. M., Lepes, D., Rubio, N., & Sheldon, K. M. (1997). Intrinsic motivation and exercise adherence. International Journal of Sport Psychology, 28(4), 335-354.

Ryan, R. M., Koestner, R., & Deci, E. L. (1991). Varied forms of persistence: When free-choice behavior is not intrinsically motivated. Motivation and Emotion, 15, 185-205.

Ryska, T. A. (2003). Enjoyment of evaluative physical activity among young participants: The role of self-handicapping and intrinsic motivation. Child Study Journal, 33(4), 213-234.

Sallis, J. F. (1994). Influences on physical activity of children, adolescents, and adults. President's Council on Physical Fitness and Sport Research Digest, 1(7).

Sallis, J. F., Prochaska, J. J., & Taylor, W. C. (2000). A review of correlates of physical activity of children and adolescents. Medicine and Science in Sport and Exercise, 32, 963 - 975.

Salmon, J., Owen, N., Crawford, D., Bauman, A., & Sallis, J. F. (2003). Physical activity and sedentary behavior: A population-based study of barriers, enjoyment, and preference. Health Psychology, 22(2), 178-188.

Scanlan, T. K., & Simons, J. P. (1992). The construct of sport enjoyment. In G. C. Roberts (Ed.), Motivation in sport and exercise (pp. 199-215). Champaign,IL: Human Kinetics.

Steiger, J. H. (1998). A note on multiple sample extensions of the RMSEA fit index. Structural Equation Modeling, 5(4), 411-419.

Stevens, M., Lemmink, K., van Heuvelen, M. J. G., de Jong, J., & Rispens, P. (2003). Groningen Active Living Model (GALM): stimulating physical activity in sedentary older adults; validation of the behavioral change model. Preventive Medicine, 37(6), 561-570.

Stevens, M., Moget, P., de Greef, M. H. G., Lemmink, K., & Rispens, P. (2000). The groningen enjoyment questionnaire: A measure of enjoyment in leisure-time physical activity. Perceptual and Motor Skills, 90(2), 601-604.

Stucky-Ropp, R. C., & DiLoenzo, T. M. (1993). Determinants of exercise in children. Preventive Medicine, 22, 880-889.

Thuot, S. M. (1995). College students' attitudes toward anticipated sport participation. Perceptual and Motor Skills, 80, 155-160.

Tsigilis, N., & Theodosiou, A. (2003). Temporal stability of the intrinsic motivation inventory. Perceptual and Motor Skills, 97(1), 271-280.

Vallerand, R. J. (2001). A hierarchical model of intrinsic and extrinsic motivation in sport and exercise. In G. C. Roberts (Ed.), Advances in motivation in sport and exercise. Champaign: Human Kinetics.

Wankel, L. M. (1993). The importance of enjoyment to adherence and psychological benefits from physical-activity. International Journal of Sport Psychology, 24(2), 151-169.

Wankel, L. M. (1997). "Strawpersons," selective reporting, and inconsistent logic: A response to Kimiecik and Harris's analysis of enjoyment. Journal of Sport & Exercise Psychology, 19(1), 98-109.

Wankel, L. M., & Kreisel, P. S. J. (1985). Factors underlying enjoyment of youth sports - Sport and age group comparisons. Journal of Sport Psychology, 7(1), 51-64.

Welk, G. J. (1999). The youth physical activity promotion model: a conceptual bridge between theory and practice. Quest, 51, 5-23.

Whitehead, J., & Corbin, C. B. (1991). Youth fitness testing: the effect of percentile-based evaluative feedback on intrinsic motivation. Research Quarterly for Exercise and Sport, 62(2), 225-231.

Wilson, P. M., Rodgers, W. M., & Fraser, S. N. (2002). Cross-validation of the revised motivation for physical activity measure in active women. Research Quarterly for Exercise and Sport, 73(4), 471-477.

Wininger, S. R. (2002). Instructors' and classroom characteristics associated with exercise enjoyment

    by females. Perceptual and Motor Skills, 94(2), 395-398.

Wininger, S. R., & Pargman, D. (2003). Assessment of factors associated with exercise enjoyment.

    Journal of Music Therapy, 40(1), 57-73.

Table 4.1. Sample size, min/max of observed scores, mean, standard deviation (SD), skewness, kurtosis, internal consistency reliabilities (α), and range of the item total correlations (ITC) for the five enjoyment scales.

	Scale (#items)	N	Min/Max	Mean	SD	Skew	Kurt	α	ITC range
1	MPAMR-joy (7)	978	7 - 49	35.311	10.099	-0.753	0.061	0.939	0.685-0.857
2	GEQ (10)	957	10 - 50	40.659	7.448	-0.989	1.280	0.919	0.627-0.770
3	PACES-M (16)	929	23 - 80	66.183	10.320	-0.852	0.653	0.929	0.494-0.757
4	IMI-joy (7)	959	7 - 49	36.114	9.215	-0.638	0.051	0.933	0.670-0.891
5	PACES-joy (10)	973	11 -70	53.997	11.100	-0.584	-0.024	0.920	0.464-0.825

Note: Skew = Skewness; Kurt = Kurtosis; ITC = Item-Total Correlation; MPAMR-joy = Motivations for physical activity measure-revised enjoyment scale; GEQ = Groningen enjoyment Questionnaire; PACES-M = Physical Activity Enjoyment Scale – modified; IMI-joy = Intrinsic motivation inventory-enjoyment/interest scale; PACES-joy = Physical Activity Enjoyment Scale – Enjoyment subscale.

Table 4.2. Fit of confirmatory factor analysis (CFA) models tested for each enjoyment scale.

		χ^2	DF	NNFI	CFI	ECVI	RMSEA (90% CI)
PACES-M							
Single Factor - CUs among positive items		275.2	68	0.953	0.977	0.434	0.055 (0.048 - 0.061)
PACES-enjoy							
Single Factor (10-items)		216.35	35	0.953	0.970	0.270	0.071 (0.062 - 0.080)
GEQ							
Single Factor		627.7	35	0.851	0.905	0.673	0.129 (0.120 - 0.138)
Single factor with CUs among items 7-10		281.68	29	0.923	0.960	0.346	0.092 (0.083 - 0.102)
Invariance	RI 1	314.8	58	0.913	0.954	0.485	0.096 (0.086 - 0.107)
	RI 2	336.346	67	0.921	0.952	0.489	0.092 (0.082 - 0.102)
	GI 1	322.013	58	0.920	0.958	0.456	0.095 (0.085 - 0.105)
	GI 2	329.969	67	0.931	0.958	0.447	0.088 (0.078 - 0.098)
MPAMR-joy							
Single Factor		165.16	14	0.949	0.974	0.203	0.103 (0.089 - 0.117)
Invariance	RI 1	202.789	28	0.936	0.968	0.303	0.115 (0.100 - 0.130)
	RI 2	207.983	34	0.948	0.968	0.296	0.105 (0.091 - 0.117)
	GI 1	194.53	28	0.943	0.972	0.273	0.107 (0.093 - 0.123)
	GI 2	203.424	34	0.952	0.971	0.27	0.099 (0.086 - 0.112)
IMI-joy							
Single Factor		475.265	14	0.847	0.924	0.506	0.180 (0.166 - 0.194)
Single factor with CU between negative items		82.01	13	0.975	0.989	0.123	0.072 (0.058 - 0.087)
Invariance	RI 1	94.945	26	0.974	0.988	0.193	0.075 (0.059 - 0.092)
	RI 2	106.969	32	0.977	0.987	0.193	0.071 (0.057 - 0.085)
	GI 1	96.485	26	0.975	0.988	0.181	0.074 (0.058 - 0.089)
	GI 2	108.255	32	0.978	0.987	0.180	0.068 (0.054 - 0.082)

Note: MPAM-R-joy = Motivations for Physical Activity Measure Enjoyment scale; GEQ = Groningen Enjoyment Questionnaire; PACES-M = Modified version of the Physical Activity Enjoyment Scale; IMI-joy = Enjoyment/Interest scale from the Intrinsic Motivation Inventory; PACES-enjoy = 10-item subscale from the original Physical Activity Enjoyment Scale; CU = Correlated uniqueness. The measures of fit include: χ^2 = Chi-square; DF = degrees of freedom; NNFI = Non-normed fit index; CFI = comparative fit index; ECVI= Expected cross-validation index; RMSEA = Root mean square error of approximation. In the invariance analysis RI = Racial invariance and GI = gender invariance. In models RI1 and GI1 the factor structure is held equivalent between groups. In models RI2 and GI2 the factor structure and factor loadings are constrained to be equivalent between groups.

Table 4.3. Factor and Bivariate correlations among enjoyment scales.

	Scale	1	2	3	4	5
			Factor Correlations			
1.	MPAMR-joy	--	0.550	0.705	0.723	0.694
2.	GEQ		--	0.739	0.726	0.590
3.	PACES-M			--	0.874	0.771
4.	IMI-joy				--	0.732
5.	PACES-joy					--
			Bivariate Correlations			
1.	MPAMR-joy	--	0.539	0.652	0.700	0.656
2.	GEQ		--	0.709	0.719	0.589
3.	PACES-M			--	0.825	0.721
4.	IMI-joy				--	0.705
5.	PACES-joy					--

Note: Factor correlations are from model M2, other correlations are Pearson product moment correlations. MPAM-R-joy = Motivations for Physical Activity Measure Enjoyment scale; GEQ = Groningen Enjoyment Questionnaire; PACES-M = Modified Physical Activity Enjoyment Scale; IMI-joy = Enjoyment/Interest scale from the Intrinsic Motivation Inventory; PACES-joy = 10-item enjoyment scale from the original Physical Activity Enjoyment Scale.

Table 4.4. Fit and comparison of confirmatory factor analysis models used to examine convergent validity and method effects.

Model	χ^2	DF	NNFI	CFI	ECVI	RMSEA (90% CI)
M1 - single enjoyment factor	11180.30	1132	0.707	0.740	11.317	0.093 (0.092 - 0.095)
M2 - 5 correlated factors	4261.07	1122	0.908	0.919	4.567	0.052 (0.051 - 0.054)
M3 - single enjoyment & 5 method factors	3662.13	1082	0.921	0.933	4.059	0.048 (0.047 - 0.050)
M4 - M3 + correlated errors among like items	3097.97	1029	0.934	0.946	3.611	0.044 (0.043 - 0.046)

Model Comparisons	$\Delta\chi^2$	ΔDF	p	ΔCFI
M1 vs. M2	6919.24	10	<0.001	-0.179
M2 vs. M3	598.94	40	<0.001	-0.014
M3 vs. M4	564.156	53	<0.001	-0.013

Note. The measures of fit include: χ^2 = Chi-square; DF = degrees of freedom; NNFI = Non-normed fit index; CFI = comparative fit index; ECVI= Expected cross-validation index; RMSEA = Root mean square error of approximation. For model comparisons Δ represents change in a given value between models.

Table 4.5. Bivariate correlations between measures of enjoyment and physical activity, perceived competence, and extrinsic motives for physical activity participation.

	Physical Activity				Perceived Competence		Extrinsic Motivation		
	Total	VIG	MOD	Work	IMI	MPAM	MPAMR Fitness	MPAMR Social	MPAMR Appear
1 MPAMR-joy	0.223	0.272	0.105	<u>0.067</u>	0.548	0.847	0.523	0.520	0.154
2 GEQ	0.249	0.289	0.115	0.102	0.608	0.537	0.387	0.285	0.151
3 PACES-M	0.229	0.284	<u>0.06</u>	<u>0.092</u>	0.659	0.605	0.442	0.226	0.148
4 IMI-joy	0.286	0.34	0.107	<u>0.077</u>	0.775	0.675	0.404	0.356	0.088
5 PACES-joy	0.219	0.271	0.081	<u>0.047</u>	0.578	0.560	0.330	0.269	<u>0.034</u>

Note: All correlations are Pearson product moment correlations. Values that are <u>underlined</u> are NOT significantly greater than zero. Physical activity estimates were calculated from responses to the International Physical Activity Questionnaire as MET-min/week: Total = Total activity; VIG = Vigorous activity; MOD = Moderate activity. Work physical activity = Work index from Baecke Questionnaire of Habitual Physical Activity. Extrinsic Motivations are from the Motivations for Physical Activity Measure - Revised. MPAM-R-joy = Motivations for Physical Activity Measure Enjoyment scale; GEQ = Groningen Enjoyment Questionnaire; PACES-M = Modified Physical Activity Enjoyment Scale; IMI-joy = Enjoyment/Interest scale from the Intrinsic Motivation Inventory; PACES-joy = 10-item enjoyment scale from the original Physical Activity Enjoyment Scale.

Table 4.6. Mean enjoyment scores and standard deviations for very low active, low active, moderately active, and highly active groups.

	Very low Active		Low active		Moderately Active		Highly Active	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
MPAMR-joy	32.349	10.778	34.155	9.306	36.243	8.605	39.844	8.472
GEQ	38.361	8.128	39.588	6.792	41.232	5.691	44.351	5.869
PACES-M	63.300	11.189	63.706	9.763	66.915	8.929	71.618	7.405
IMI-joy	32.821	9.602	34.496*	8.653	37.676*	6.985	41.262	7.109
PACES-joy	51.125	11.855	51.699*	10.731	55.500*	8.299	59.156	8.977

Note: Based on the ANOVA results, the following are true for all measures of enjoyment. Low active = Very low active; High active > all other categories; Moderately active > very low active. The asterisk (*) denotes significant differences ($p < 0.05$) between low active and moderately active groups only.
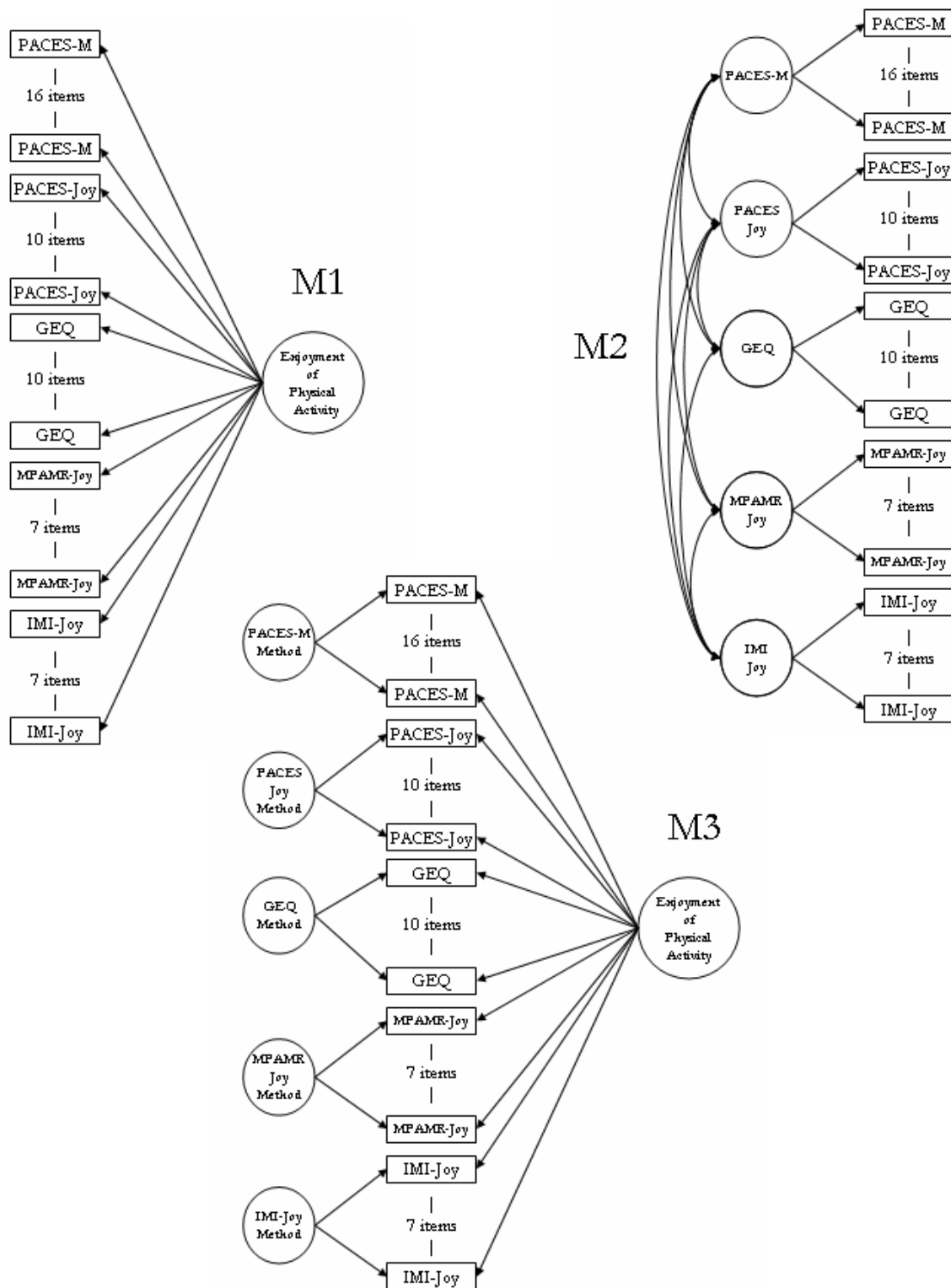
Figure 4.1. Confirmatory factor analysis models used to examine convergent validity and method effects for the five physical activity enjoyment scales. Model 1 (M1) includes a single enjoyment factor indicated by all 50 items. Model 2 (M2) includes five correlated factors. Each method/trait factor is indicated by the items from one of the enjoyment scales. Model 3 (M3) includes a single enjoyment factor indicated by all 50 items and five method factors indicated by the items from one of the enjoyment scales. For simplicity, the item uniqueness, or measure errors, and correlations between item uniquenesses are not pictured in the models.
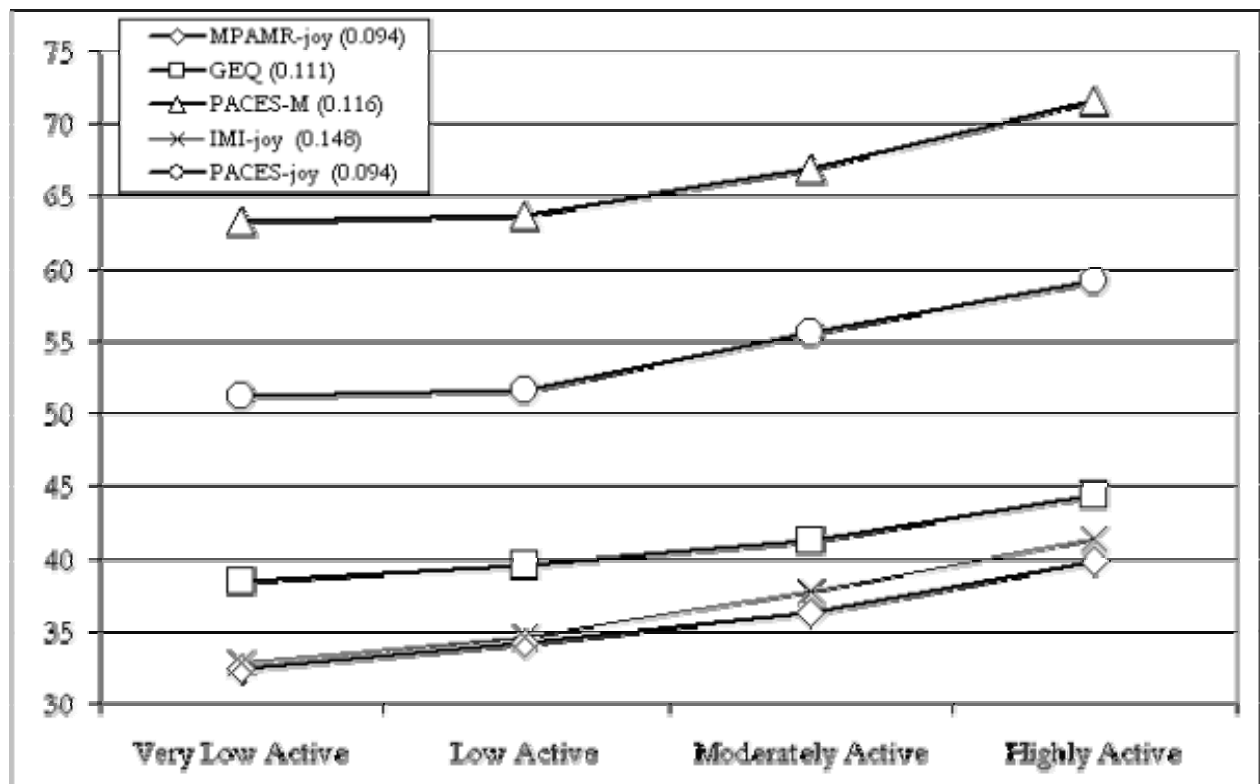
111

Figure 4.2. Plot of mean enjoyment scores for very low, low, moderately, and highly active participants. Numbers in the legend are adjusted R^2 values for the regression model predicting physical activity level from enjoyment score.

CHAPTER 5

SUMMARY AND CONCLUSIONS

The purpose of this research was to compare and extend the construct validity evidence for several measures of physical activity enjoyment. This was accomplished by: (1) comparing item content across the enjoyment scales, (2) testing the factor structure of each scale, (3) testing measurement invariance across gender and race for each scale, (4) assessing convergent validity evidence, and (5) estimating and comparing the relationship of each enjoyment scale with measures of physical activity, perceived competence, and extrinsic participation motives.

Before a scale is used to indicate a latent construct, it is important to establish that the hypothesized relationships among the items on the scale are supported. Confirmatory factor analysis can be used to test a priori hypotheses concerning relationships among indicators of a construct (i.e. the factor structure of the scale). This method can also be used to test the equivalence of the factor structure and parameter estimates across groups (i.e. factor invariance). If the factor structure and loadings of a scale are invariant between groups, it is assumed that the same construct is being measured.

Participants (N = 1023; mean age 19.60 ± 1.55 years) were recruited from basic physical education classes and the psychology research pool at two universities. Data were collected by web-based survey. Each participant completed the Physical Activity Enjoyment Scale (PACES), a modified version of the PACES (PACES-M), the interest/enjoyment (IMI-joy) and perceived competence subscales from the Intrinsic Motivation Inventory (IMI), the Revised Motivation for Physical Activity Measure (MPAMR), the Groningen Enjoyment Questionnaire (GEQ), the International Physical Activity Questionnaire, the Lipid Research Clinic Physical Activity Questionnaire, and the work index from the Baecke Questionnaire of Habitual Physical Activity.

113

The physical activity enjoyment scales compared in this study were chosen based on their use in the physical activity literature, current state of development, and item content.

A simple single factor model did not adequately account for the covariance among the items on four of the five scales. The PACES-M, GEQ, and IMI-joy models required correlated uniquenesses to improve model fit. Based on exploratory analyses, a two factor model was found to represent the data from the PACES. The first factor, PACES-joy, included items reflecting enjoyment, interest, fun, and liking, while the other factor, PACES-energy, included items related to energy and achievement. This two-factor model suggests that either enjoyment is a multidimensional construct or that the PACES measures two related, but distinct latent factors. The fit of the simple single factor model for the MPAMR-joy was acceptable, but the RMSEA was 0.10. The PACES-enjoy scale was judged to have the strongest evidence for factor validity. This was based on the fact that the single–factor model provided a good representation of the covariance among the items and did not include correlated uniquenesses, making it easier to interpret.

Establishing measurement invariance is an important piece of validity evidence supporting the use and comparison of scores between groups. These results support the conclusion that the relationship among items on the PACES, PACES-M, GEQ, IMI-joy, and MPAMR-joy are nearly equivalent between male and female and between black and white young adults.

Validity evidence supporting the convergence of measures of the same construct is important when different scales are frequently used to represent a construct. If two measures of the same trait do not correlate highly the validity of both may be called into question. Comparisons among the five enjoyment scales supported similarity of measurement but not complete convergence. Results from the confirmatory factor analysis and the correlational analyses suggested that a significant method effect was present. In most cases, the method effect seems to be an artifact that does not change the relationship of enjoyment with other measures. However, the relationship

between enjoyment and perceived competence is substantially higher when the two constructs are measured with subscales of the same instrument (i.e. IMI or MPAMR). This effect should be considered when selecting a measure of enjoyment and when testing theoretical relationships among variables measured with the same scale.

Finally, correlations were estimated between each measure of enjoyment and physical activity, perceived competence, and extrinsic motives for physical activity participation. Based on evidence from the physical activity, exercise, sport, motivation, and enjoyment literature, several hypotheses were developed with respect to the construct validity evidence for the enjoyment scores. First, it was hypothesized that the magnitude of the correlations between enjoyment and total physical activity would be 0.20 – 0.40 and that the relationship between enjoyment and work-related physical activity would be very small. Also, based on cognitive evaluation theory, it was predicted that enjoyment would exhibit a stronger relationship with an intrinsic motive (i.e. perceived competence) than with more extrinsic motives (i.e. social, appearance, and fitness). Results again supported the comparability of enjoyment measures. The correlations with various estimates of physical activity were very similar across measures of enjoyment and supported the a priori hypotheses. Relationships with perceived competence and extrinsic motives were also as hypothesized, but were larger between constructs measured with subscales of the same instrument (e.g. IMI and MPAMR).

In the future researchers should attempt to confirm and extend the findings from this study in several ways. First, although invariance was supported across race, generalization may be limited due to the small number of black males (n = 38) in the sample. Attempts should be made to cross-validate these results with larger samples of black males and to extend these findings to other racial, ethnic, and cultural groups. In addition, it will be important to determine if the scales are invariant across a wide range of ages. If they are invariant, scores could be used to compare enjoyment of

115

physical activity across the lifespan and to compare changes in enjoyment over longer periods of time. Second, no direct comparisons have been made across different types of activity. Changing the context (i.e. sports, fitness activities, vigorous activities, sedentary activities, etc.) may affect the measurement structure of the scale.

The evidence also suggests a need to reevaluate items on the enjoyment scales. This line of research could explore a number of issues related to item function and content. For example, the dimensionality of the enjoyment construct could be investigated further. Young adults in this sample made some distinction between general positive affect and feelings of achievement and vitality associated with physical activity. It is not clear whether these factors represent two different constructs or two facets of enjoyment. Also, it seems necessary to determine the most economical measure that fully represents the enjoyment construct. In this study, the length of the scale did not affect the findings. If a three- or four-item scale is found to be equivalent to a 10- or 18-item scale, use of the short scale would be preferred in a number of situations. Finally, the large percentage of people scoring above the midpoint for each enjoyment scale along with the fact that mean enjoyment scores did not differ between very low and low active individuals suggests a floor effect for these measures of enjoyment. If new items are developed an attempt should be made to create items that discriminate better across all levels of physical activity.

In conclusion, scores from the five measures of enjoyment are comparable. No single instrument stood out as a "best" measure of enjoyment, but the 10-item PACES-enjoy scale has a sound factor structure with no correlated uniquenesses. Slight differences in factor structure and the presence of method effects do not affect the relationship of any enjoyment scale with physical activity. The method effect does, however, inflate the relationship between constructs measured with subscales from the same instrument (i.e IMI and MPAMR). The measurement properties of the scales are the same for males and females and for black and white young adults. The relationship of

enjoyment and intrinsic and extrinsic motives is consistent across scales, strengthening the nomological network for the enjoyment construct. Overall, this research supports the construct validity and similarity of scores from the PACES, PACES-M, GEQ, IMI-joy, and MPAMR-joy.

REFERENCES

Ainsworth, B. E., Jacobs, D. R., & Leon, A. S. (1993). Validity and reliability of self-reported

    physical-activity status - The lipid-research-clinics questionnaire. Medicine and Science in

    Sports and Exercise, 25, 92-98.

Ajzen, I. (1988). Attitudes, personality and behavior. Chicago: Dorsey Press.

Ajzen, I., & Fishbein, M. (1974). Factors influencing intentions and the intention-behavior relation.

    Human Relations, 27 (1), 1-15.

Akaike, H. (1987). Factor analysis and AIC. Psychometrica, 52, 317-332.

Albanes, D., Conway, J. M., Taylor, P. R., Moe, P. W., & Judd, J. (1990). Validation and comparison

    of eight physical activity questionnaires. Epidemiology, 1, 65-71.

American Educational Research Association (1999). Standards for educational and psychological

    testing. Washington, DC: AERA.

Arbuckle, J. L. (1996). Full information estimation in the presence of incomplete data. In G. A.

    Marcoulides & R. E. Schumacker (Eds.), Advanced structural equation modeling: Issues and

    techniques (pp. 243-277). Mahwah, NJ: Lawrence Erlbaum.

Arbuckle, J. L. (2003). Amos 5.0 update to the Amos user's guide. Chicago, IL: SPSS

Arbuckle, J. L., & Wothke, W. (1999). Amos 4.0 user's guide. Chicago, Il: SmallWaters Corporation.

Baecke, J. A. H., Burema, J., & Frijters, J. E. R. (1982). A short questionnaire for the measurement

    of habitual physical activity in epidemiological studies. American Journal of Clinical

    Nutrition, 36, 936-942.

Bandalos, D. L. (1993). Factors influencing cross-validation of confirmatory factor-analysis models.

    Multivariate Behavioral Research, 28(3), 351-374.

Bandura, A. (1986). Social foundations of thought and action. Englewood Cliffs, NJ: Prentice-Hall.

Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance-structures. Psychological Bulletin, 88(3), 588-606.

Blair, S., Kohl, H., Paffenbarger, R., Clark, D., Cooper, K., & Gibbons, L. (1989). Physical fitness and all-cause mortality: A prospective study of healthy men and women. Journal of the American Medical Association, 262, 2395-2401.

Bollen, K. A. (1989). Structural equations with latent variables. New York: John Wiley.

Boomsma, A. (2000). Reporting analyses of covariance structures. Structural Equation Modeling, 7(3), 461-483.

Browne, M. W., & Cudeck, R. (1989). Single sample cross-validation indexes for covariance-structures. Multivariate Behavioral Research, 24(4), 445-455.

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), Testing structural equation models (pp. 136-162). Newbury Park, CA: Sage Publications, Inc.

Bungum, T., Dowda, M., Weston, A., Trost, S. G., & Pate, R. R. (2000). Correlates of physical activity in male and female youth. Pediatric Exercise Science, 12(1), 71-79.

Byrne, B. M. (1998). Structural equation modeling with LISREL, PRELIS, and SIMPLIS: Basic concepts, applications, and programming. Mahwah, N.J.: Lawrence Erlbaum.

Calfas, K. J., Sallis, J. F., Nichols, J. F., Sarkin, J. A., Johnson, M. F., Caparosa, S., (2000). Project GRAD: Two-year outcomes of a randomized controlled physical activity intervention among young adults. American Journal of Preventive Medicine, 18(1), 28-37.

Campbell, D.T. & Fiske, D.W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. Psychological Bulletin, 56, 81-105.

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. Structural Equation Modeling, 9(2), 233-255.

Cooper Institute/The American College of Sports Medicine: 1997 Physical Activity Interventions Conference. (1998). American Journal of Preventive Medicine, 15 (4).

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

Crocker, P. R. E., Bouffard, M., & Gessaroli, M. E. (1995). Measuring enjoyment in youth sport settings - A confirmatory factor-analysis of the physical-activity enjoyment scale. Journal of Sport and Exercise Psychology, 17(2), 200-205.

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

Cronbach, L.J. (1988). Five perspectives on the validity argument. In H. Waner & H.I. Braun (Eds.), Test validity (pp. 3-17). Hillsdale, NJ: Erlbaum.

Csikszentmihalyi, M. (1975). Beyond boredom and anxiety: Experience flow in work and play. San Francisco: Jossey-Bass.

Csikszentmihalyi, M. (1990). Flow: the psychology of optimal experience. New York: Harper and Row.

Davison, K. K., & Birch, L. L. (2001). Child and parent characteristics as predictors of change in girls' body mass index. International Journal of Obesity, 25(12), 1834-1842.

De Bourdeaudhuij, I., & Sallis, J. F. (2002). Relative contribution of psychosocial variables to the explanation of physical activity in three population-based adult samples. Preventive Medicine, 34, 279-288.

de Charms, R. (1968). Personal causation: the internal affective determinants of behavior. New York: Irvington.

Deci, E. L. (1975). Intrinsic motivation. New York: Plenum.

Deci, E. L., & Ryan, R. M. (1985). <u>Intrinsic motivation and self-determination in human behavior.</u> New York: Plenum.

Dergance, J. M., Calmbach, W. L., Dhanda, R., Miles, T. P., Hazuda, H. P., & Mouton, C. P. (2003). Barriers to and benefits of leisure time physical activity in the elderly: Differences across cultures. <u>Journal of the American Geriatrics Society, 51</u>(6), 863-868.

DiLorenzo, T. M., Stucky-Ropp, R. C., Vander Wal, J. S., & Gotham, H. J. (1998). Determinants of exercise among children. II. A longitudinal analysis. <u>Preventive Medicine, 27</u>(3), 470-477.

Dishman, R. K., Motl, R. W., Saunders, R., Felton, G., Ward, D. S., Dowda, M. (2005). Enjoyment mediates effects of a school-based physical-activity intervention. <u>Medicine and Science in Sports and Exercise, 37</u>(3), 478-487.

Dishman, R. K., & Sallis, J. F. (1994). Determinants and interventions for physical activity and exercise. In C. Bouchard, R. J. Shephard & T. Stephens (Eds.), <u>Physical activity, fitness, and health: International proceedings and consensus statement </u>(pp. 214 - 238). Champaign, IL: Human Kinetics.

Dunn, A. L., Trivedi, M. H., & O'Neal, H. A. (2001). Physical activity dose-response effects on outcomes of depression and anxiety. <u>Medicine and Science in Sport and Exercise, 33</u>, s587-s597.

Ebbeck, V., Gibbons, S. L., & Loken-Dahle, L. J. (1995). Reasons for adults participation in physical: an interactional approach. <u>International Journal of Sport Psychology, 26</u>, 262-275.

Enders, C. K., & Bandalos, D. L. (2001). The relative performance of full information maximum likelihood estimation for missing data in structural equation models. <u>Structural Equation Modeling, 8</u>, 430-457.

Fischer, D. G., & Fick, C. (1993). Measuring social desirability: short forms of the Marlowe-Crowne social desirability scale. <u>Educational and Psychological Measurement, 53,</u> 417 - 424.

Frederick, C. M., Morrison, C., & Manning, T. (1996). Motivation to participate, exercise affect, and outcome behaviors toward physical activity. Perceptual and Motor Skills, 82(2), 691-701.

Frederick, C. M., & Ryan, R. M. (1993). Differences in motivation for sport and exercise and their relations with participation and mental health. Journal of Sports Behavior, 16, 124-146.

Frederick-Recascino, C. M., & Schuster-Smith, H. (2003). Competition and intrinsic motivation in physical activity: a comparison of two groups. Journal of Sport Behavior, 26(3), 240-254.

Garcia, A. W., & King, A. C. (1991). Predicting long-term adherence to aerobic exercise: A comparison of two models. Journal of Sport and Exercise Psychology, 13, 394 - 410.

Gill, D., Gross, J., & Huddleston, S. (1983). Participation motivation in youth sports. International Journal of Sports Psychology, 14(1), 11-14.

Gould, D., & Horn, T. (1984). Participation motivation in young athletes. In (Eds.),. In J. M. Silva & R. S. Weinberg (Eds.), Psychological foundations of sport (pp. 359-370). Champaign, IL: Human Kinetics.

Haase, A., Steptoe, A., Sallis, J. F., & Wardle, J. (2004). Leisure-time physical activity in university students from 23 countries: Associations with health beliefs, risk awareness, and national economic development. Preventive Medicine, 39, 182-190.

Hagger, M.S., Chatzisarantis, N. & Biddle, S.J.H. (2002). A meta-analytic review of the theories of reasoned action and planned behavior in physical activity: Predictive validity and the contribution of additional variables. Journal of Sport and Exercise Psychology, 24 (1), 3-32.

Hales, D. P. (2005). Factor validity, invariance, and comparison of several measures of physical activity enjoyment. Unpublished Dissertation, University of Georgia, Athens.

Heck, T. A., & Kimiecik, J. C. (1993). What is exercise enjoyment?: A qualitative investigation of adult exercise maintainers. Wellness Perspectives: Research, Theory and Practice, 10(1).

Helmrich, S. P., Ragland, D. R., Leung, R. W., & Paffenbarger, R. (1991). Physical activity and reduced occurrence of non-insulin-dependent diabetes mellitus. New England Journal of Medicine, 325(147-152).

Horn, J. L., & McArdle, J. J. (1992). A practical and theoretical guide to measurement invariance in aging research. Experimental Aging Research, 18(3-4), 117-144.

Hu, L. T., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. Psychological Methods, 3(4), 424-453.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. Structural Equation Modeling, 6(1).

Iso-Ahola, S. E., & St. Clair, B. (2000). Toward a theory of exercise motivation. Quest, 52, 131-147.

Izard, C. E. (1977). Human emotions. New York: Plenum.

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

Jagacinski, C. M., & Duda, J. L. (2001). A comparative analysis of contemporary achievement goal orientation measures. Educational and Psychological Measurement, 61(6), 1013-1039.

Johnson, N. A., & Heller, R. F. (1998). Prediction of patient nonadherence with home-based exercise for cardiac rehabilitation: The role of perceived barriers and perceived benefits. Preventive Medicine, 27(1), 56-64.

Joreskog, K. G. (1993). Testing structural equation models. In K. A. Bollen & J. S. Long (Eds.), Testing structural equation models. Newbury Park, CA: Sage.

Jöreskog, K. G., & Sörbom, D. (1996a). LISREL 8: User's reference guide. Chicago, IL: Scientific Software International.

Joreskog, K., & Sorbom, D. (1996b). <u>PRELIS 2: Users reference guide</u>. Hillsdale, NJ: Lawrence Erlbaum.

Kafatos, A., Manios, Y., Markatji, I., Giachetti, I., Vaz de Almeida, M. D., & Engstrom, L. M. (1999). Regional, demographic and national influences on attitudes and beliefs with regard to physical activity, body weight and health in a nationally representative sample in the european union. <u>Public Health Nutrition, 2,</u> 87-95.

Kendzierski, D., & Decarlo, K. J. (1991). Physical activity enjoyment scale - two validation studies. <u>Journal of Sport and Exercise Psychology, 13</u>(1), 50-64.

Kenny, D. A., & Kashy, D. A. (1992). Analysis of the multitrait-multimethod matrix by confirmatory factor analysis. <u>Psychological Bulletin, 112</u>(1), 165-172.

Kenny, D. A., & McCoach, D. B. (2003). Effect of the number of variables on measures of fit in structural equation modeling. <u>Structural Equation Modeling, 10</u>, 333-351.

Kimiecik, J. C., & Harris, A. T. (1996). What is enjoyment? A conceptual/definitional analysis with implications for sport and exercise psychology. <u>Journal of Sport and Exercise Psychology, 18</u>(3), 247-263.

Kjelsas, E., & Augestad, L. B. (2003). Gender differences in competitive runners and their motive for physical activity. <u>European Journal of Psychiatry, 17</u>(3), 157-171.

Klentrou, P., Hay, J., & Plyley, M. (2003). Habitual physical activity levels and health outcomes of Ontario youth. <u>European Journal of Applied Physiology, 89</u>(5), 460-465.

Kline, R. B. (1998). <u>Principles and practices of structural equation modeling</u>. New York: The Guilford Press.

Knowler, W. C., Barrett-Connor, E., Fowler, S. E., Hamman, R. F., Lachin, J. M., Walker, E. A. (2002). Reduction in incidence of type 2 diabetes with lifestyle intervention or metformin. <u>New England Journal of Medicine, 346</u>(6), 393-403.

Lance, C. E., Noble, C. L., & Scullen, S. E. (2002). A critique of the correlated trait-correlated method and correlated uniqueness models for multitrait-multimethod data. Psychological Methods, 7(2), 228-244.

Lance, C. E., & Vandenberg, R. J. (2001). Confirmatory factor analysis. In F. Drasgow & N. Schmitt (Eds.), Measuring and analyzing behavior in organizations (pp. 221-254). San Francisco: Jossey-Bass.

Leslie, E., Owen, N., Salmon, J., Bauman, A., Sallis, J. F., & Lo, S. K. (1999). Insufficiently active Australian college students: Perceived personal, social, and environmental influences. Preventive Medicine, 28(1), 20-27.

Lewis, B. A., Marcus, B. H., Pate, R. R., & Dunn, A. L. (2002). Psychosocial mediators of physical activity behavior among adults and children. American Journal of Preventive Medicine, 23(2), 26-35.

Long, J. S. (1983). Covariance structure models: an introduction to LISREL (Vol. 07-034). Newbury Park: Sage.

MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. Psychological Methods, 1, 130-149.

Manson, J. E., Hu, F., Rich-Edwards, J., Colditz, G., Stampfer, W., & Willet, W. (1999). A prospective study of walking as compared with vigorous exercise in the prevention of coronary heart disease in women. New England Journal of Medicine, 341(9), 650-658.

Mardia, K. V. (1970). Measures of multivariate skewness and kurtosis with applications. Biometrika, 57, 519-530.

Markland, D., & Hardy, L. (1997). On the factorial and construct validity of the intrinsic motivation inventory: Conceptual and operational concerns. Research Quarterly for Exercise and Sport, 68(1), 20-32.

Marsh, H. W. (1996). Positive and negative global self-esteem: A substantively meaningful distinction or artifactors? Journal of Personality and Social Psychology, 70(4), 810-819.

Marsh, H. W., Hau, K., & Wen, Z. (2004). In search of golden rules: comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. Structural Equation Modeling, 11(3), 320-341.

Marsh, H. W., Richards, G. E., Johnson, S., Lawrence, R., & Tremayne, P. (1994). Physical self-description questionnaire: psychometric properties and a multitrait-multimethod analysis of relations to existing instruments. Journal of Sport and Exercise Psychology, 16, 270-305.

Matsumoto, H., & Takenaka, K. (2004). Motivational profiles and stages of exercise behavior change. International Journal of Sport and Health Science, 2, 89-96.

McAuley, E., Duncan, T., & Tammen, V. V. (1989). Psychometric properties of the intrinsic motivation inventory in a competitive sport setting - A confirmatory factor-analysis. Research Quarterly for Exercise and Sport, 60, 48-58.

McAuley, E., Jerome, G. J., Elavsky, S., Marquez, D. X., & Ramsey, S. N. (2003). Predicting long-term maintenance of physical activity in older adults. Preventive Medicine, 37(2), 110-118.

McAuley, E., Wraith, S., & Duncan, T. E. (1991). Self-efficacy, perceptions of success, and intrinsic motivation for exercise. Journal of Applied Social Psychology, 21, 139-155.

Messick, S. (1975). The standard problem: meaning and values in measurement and evaluation. American Psychologist, 30, 955-966.

Messick, S. (1989). Validity. In R. L. Linn (Ed.), Educational measurement (third ed., pp. 13-103). New York: Macmillan.

Messick, S. (1995). Validity of psychological-assessment - Validation of inferences from persons responses and performances as scientific inquiry into score meaning. American Psychologist, 50(9), 741-749.

Moore, J. B., Yin, Z., & Gutin, B. (2004). Enjoyment of physical activity in children: Instrument validation in the MCG FitKid project. Journal of Sport and Exercise Psychology, 26, S138-S139.

Motl, R. W., Berger, B. G., & Leuschen, P. S. (2000). The role of enjoyment in the exercise-mood relationship. International Journal of Sport Psychology, 31(3), 347-363.

Motl, R. W., Dishman, R. K., Saunders, R., Dowda, M., Felton, G., & Pate, R. R. (2001). Measuring enjoyment of physical activity in adolescent girls. American Journal of Preventive Medicine, 21(2), 110-117.

Mullan, E., & Markland, D. (1997). Variations in self-determination across the stages for exercise in adults. Motivation and Emotion, 21(4), 349-362.

Neumark-Sztainer, D., Story, M., Hannan, P. J., Tharp, T., & Rex, J. (2003). Factors associated with changes in physical activity - A cohort study of inactive adolescent girls. Archives of Pediatrics and Adolescent Medicine, 157(8), 803-810.

Nies, M. A., Vollman, M., & Cook, T. (1999). African american women's experiences with physical activity in their daily lives. Public Health Nursing, 16(1), 23-31.

Nigg, C. R., Allegrante, J. P., & Ory, M. (2002). Theory-comparison and multiple-behavior research: common themes advancing health behavior research. Health Education Research, 17(5), 670-679.

Paffenbarger, R., Wing, A., Hyde, R., & Jung, D. (1983). Physical activity and incidence of hypertension in college alumni. American Journal of Epidemiology, 117, 245-257.

Petri, H. L., & Govern, J. M. (2004). Motivation: theory, research, and applications (fifth ed.). Belmont, CA: Wadsworth/Thomson Learning.

Prochaska, J. J., Sallis, J. F., Slymen, D. J., & McKenzie, T. L. (2003). A longitudinal study of children's enjoyment of physical education. Pediatric Exercise Science, 15(2), 170-178.

Reiss, S. (2004). Multifaceted nature of intrinsic motivation the theory of 16 basic desires. Review of General Psychology, 8(3), 179-193.

Rovniak, L. S., Anderson, E. S., Winett, R. A., & Stephens, R. S. (2002). Social cognitive determinants of physical activity in young adults: A prospective structural equation analysis. Annals of Behavioral Medicine, 24(2), 149-156.

Ryan, R. M. (1982). Control and information in the intrapersonal sphere: An extension of cognitive evaluation theory. Journal of Personality and Social Psychology, 43, 450-461.

Ryan, R. M., & Deci, E. L. (2000a). Intrinsic and extrinsic motivations: Classic definitions and new directions. Contemporary Educational Psychology, 25, 54-67.

Ryan, R. M., & Deci, E. L. (2000b). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. American Psychologist, 55, 68-78.

Ryan, R. M., Frederick, C. M., Lepes, D., Rubio, N., & Sheldon, K. M. (1997). Intrinsic motivation and exercise adherence. International Journal of Sport Psychology, 28(4), 335-354.

Ryan, R. M., Koestner, R., & Deci, E. L. (1991). Varied forms of persistence: When free-choice behavior is not intrinsically motivated. Motivation and Emotion, 15, 185-205.

Ryska, T. A. (2003). Enjoyment of evaluative physical activity among young participants: the role of self-handicapping and intrinsic motivation. Child Study Journal, 33(4), 213-234.

Saelens, B. E., & Epstein, L. H. (1999). The rate of sedentary activities determines the reinforcing value of physical activity. Health Psychology, 18(6), 655-659.

Salguero, A., Gonzalez-Boto, R., Tuero, C., & Marquez, S. (2003). Identification of dropout reasons in young competitive swimmers. Journal of Sports Medicine and Physical Fitness, 43(4), 530-534.

Sallis, J. F. (1994). Influences on physical activity of children, adolescents, and adults. President's Council on Physical Fitness and Sport Research Digest, 1(7).

Sallis, J. F., Alcaraz, J. E., McKenzie, T. L., & Hovell, M. F. (1999). Predictors of change in

    children's physical activity over 20 months: variations by gender and level of adiposity.

    <u>American Journal of Preventive Medicine, 16</u>(3), 222-229.

Sallis, J. F., Prochaska, J. J., & Taylor, W. C. (2000). A review of correlates of physical activity of

    children and adolescents. <u>Medicine and Science in Sport and Exercise, 32</u>, 963 - 975.

Sallis, J. F., Prochaska, J. J., Taylor, W. C., Hill, J. O., & Geraci, J. C. (1999). Correlates of physical

    activity in a national sample of girls and boys in grades 4 through 12. <u>Health Psychology,</u>

    <u>18</u>(4), 410-415.

Salmon, J., Owen, N., Crawford, D., Bauman, A., & Sallis, J. F. (2003). Physical activity and

    sedentary behavior: A population-based study of barriers, enjoyment, and preference. <u>Health</u>

    <u>Psychology, 22</u>(2), 178-188.

Scanlan, T. K., Carpenter, P. J., Schmidt, G. W., Simons, J. P., & Keeler, B. (1993). An introduction

    to the sport commitment model. <u>Journal of Sport and Exercise Psychology, 15</u>(1), 1-15.

Scanlan, T. K., & Lewthwaite, R. (1986). Social psychological-Aspects of competition for male youth

    sport participants .4. Predictors of enjoyment. <u>Journal of Sport Psychology, 8</u>(1), 25-35.

Scanlan, T. K., & Simons, J. P. (1992). The construct of sport enjoyment. In G. C. Roberts (Ed.),

    <u>Motivation in sport and exercise</u> (pp. 199-215). Champaign,IL: Human Kinetics.

Scanlan, T. K., Stein, G. L., & Ravizza, K. (1989). An in-depth study of former elite figure skaters:

    II. sources of enjoyment. <u>Journal of Sport and Exercise Psychology, 11</u>, 65-83.

Schleppe, S. (1993). <u>The physical activity enjoyment scale: Reliability and validity.</u> University of

    Alberta (Canada), Canada.

Steiger, J. H. (1998). A note on multiple sample extensions of the RMSEA fit index. <u>Structural</u>

    <u>Equation Modeling, 5</u>(4), 411-419.

Steiger, J. H., & Lind, J. C. (1980). Statistically based tests for the number of common factors. Paper presented at the Psychometric Society Annual Meeting, Iowa City, IA.

Stein, G. L., Kimiecik, J. C., Daniels, J., & Jackson, S. A. (1995). Psychological antecedents of flow in recreational sport. Personality and Social Psychology Bulletin, 21(2), 125-135.

Stevens, M., Lemmink, F., de Greef, N. H. G., & Rispens, P. (2000). Groningen Active Living Model (GALM): Stimulating physical activity in sedentary older adults; First results. Preventive Medicine, 31(5), 547-553.

Stevens, M., Lemmink, K., van Heuvelen, M. J. G., de Jong, J., & Rispens, P. (2003). Groningen Active Living Model (GALM): stimulating physical activity in sedentary older adults; validation of the behavioral change model. Preventive Medicine, 37(6), 561-570.

Stevens, M., Moget, P., de Greef, M. H. G., Lemmink, K., & Rispens, P. (2000). The Groningen Enjoyment Questionnaire: A measure of enjoyment in leisure-time physical activity. Perceptual and Motor Skills, 90(2), 601-604.

Stucky-Ropp, R. C., & DiLoenzo, T. M. (1993). Determinants of exercise in children. Preventive Medicine, 22, 880-889.

Thuot, S. M. (1995). College students' attitudes toward anticipated sport participation. Perceptual and Motor Skills, 80, 155-160.

Trost, S. G., Pate, R. R., Saunders, R., Ward, D. S., Dowda, M., & Felton, G. (1997). A prospective study of the determinants of physical activity in rural fifth-grade children. Preventive Medicine, 26(2), 257-263.

Tsigilis, N., & Theodosiou, A. (2003). Temporal stability of the intrinsic motivation inventory. Perceptual and Motor Skills, 97(1), 271-280.

130

US Department of Health and Human Services (1996). <u>Physical activity and health: A report of the surgeon general</u>. Atlanta, GA: Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion.

US Department of Health and Human Services (2003). <u>Behavior risk factor surveillance system</u>. Atlanta, GA:, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion.

US Department of Health and Human Services (2003). <u>Youth risk behavior survey</u>. Atlanta, GA:, Centers for Disease Control and Prevention, Chronic Disease and Health Promotion.

U.S. Department of Health and Human Services (2004). <u>National health interview survey</u>. Washington, D.C.:, Public Health Service, Centers for Disease Control and Prevention, National Center for Health Statistics.

Vallerand, R. J. (2001). A hierarchical model of intrinsic and extrinsic motivation in sport and exercise. In G. C. Roberts (Ed.), <u>Advances in Motivation in Sport and Exercise</u>. Champaign: Human Kinetics.

Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. <u>Organizational Research Methods, 3</u>(1), 4-70.

Vilhjalmsson, R., & Thorlindsson, T. (1998). Factors related to physical activity: A study of adolescents. <u>Social Science and Medicine, 47</u>(5), 665-675.

Wankel, L. M. (1993). The importance of enjoyment to adherence and psychological benefits from physical-activity. <u>International Journal of Sport Psychology, 24</u>(2), 151-169.

Wankel, L. M. (1997). "Strawpersons," selective reporting, and inconsistent logic: A response to Kimiecik and Harris's analysis of enjoyment. <u>Journal of Sport and Exercise Psychology, 19</u>(1), 98-109.

Wankel, L. M., & Kreisel, P. S. J. (1985). Factors underlying enjoyment of youth sports - sport and age group comparisons. <u>Journal of Sport Psychology, 7</u>(1), 51-64.

Welk, G. J. (1999). The youth physical activity promotion model: A conceptual bridge between theory and practice. <u>Quest, 51</u>, 5-23.

White, R. (1959). Motivation reconsidered: The concept of competence. <u>Psychological Review, 66</u>, 297 - 333.

Whitehead, J. (1993). Physical activity and intrinsic motivation. <u>PCPFS Research Digest</u>, Series 1(Number 2), 1-8.

Whitehead, J., & Corbin, C. B. (1991). Youth fitness testing: the effect of percentile-based evaluative feedback on intrinsic motivation. <u>Research Quarterly for Exercise and Sport, 62</u>(2), 225-231.

Widaman, K. F. (1985). Hierarchically nested covariance structure models for multitrait-multimethod data. <u>Applied Psychological Measurement, 9</u>(1), 1-26.

Wilcox, S. (2002). Physical activity in older women of color. <u>Topics in Geriatric Rehabilitation, 18</u>(1), 21-33.

Wilcox, S., Richter, D. L., Henderson, K. A., Greaney, M. L., & Ainsworth, B. E. (2002). Perceptions of physical activity and personal barriers and enablers in african-american women. <u>Ethnicity and Disease, 12</u>(3), 316-319.

Williamson, D. F. (1993). Descriptive epidemiology of body weight changes in U.S. adults. <u>Annals of Internal Medicine, 119</u>, 649-649.

Wilson, P. M., Rodgers, W. M., & Fraser, S. N. (2002). Examining the psychometric properties of the behavioral regulation in exercise questionnaire. <u>Measurement in Physical Education and Exercise Science, 6</u>(1), 1-21.

Wininger, S. R. (2002). Instructors' and classroom characteristics associated with exercise enjoyment by females. <u>Perceptual and Motor Skills, 94</u>(2), 395-398.

Wininger, S. R., & Pargman, D. (2003). Assessment of factors associated with exercise enjoyment. Journal of Music Therapy, 40(1), 57-73.

Wood, T. M. (1989). The changing nature of norm-referenced validity. In M. J. Safrit & T. M. Wood (Eds.), Measurement concepts in physical education and exercise science (pp. 23-44). Champaign, Il: Human Kinetics.

Young, D. R., Gittelsohn, J., Charleston, J., Felix-Aaron, K., & Appel, L. J. (2001). Motivations for exercise and weight loss among African-American women: focus group results and their contribution towards program development. Ethnicity and Health, 6, 227-245.

Zook, A., & Sipps, G. J. (1985). Cross-validation of a short form of the Marlowe-Crowne social desirablity scale. Journal of Clinical Psychology, 41, 236-238.