THE CREATION AND VALIDATION OF THE PHYSICAL EDUCATOR EFFICACY SCALE FOR TEACHING LIFETIME PHYSICAL ACTIVITIES

A Dissertation

Presented to

The Faculty of the Curry School of Education

University of Virginia

In Partial Fulfillment

Of the Requirements for the Degree

Doctor of Philosophy

---

by

Kason M. O'Neil, B.S.Ed., M.S.S.

June, 2014

ABSTRACT

Due to the emphasis numerous national, state, and local organizations have placed on the promotion of lifetime physical activity in our school physical education programs, it is imperative that researchers continue to examine how lifetime physical activities are being taught in schools, and what teachers specifically think about their confidence towards instruction of these activities (AAHPERD 2103; Fairclough, Stratton, & Baldwin, 2002; NASPE 2007; 2008; 2009). As a result, a valid measure is needed to assess how teachers perceive their abilities to teach lifetime physical activities. Currently, there is no existing valid instrument that measures self-efficacy perceptions of physical educators towards the instruction of lifetime physical activities. To address the gap in the research, a new scale with strong evidence for validity and reliability is needed.

The purpose of this study was to develop and begin the validation process of an instrument that measures efficacy perceptions of physical educators towards teaching lifetime physical activities. This instrument, the *Physical Educator Efficacy Scale for Teaching Lifetime Physical Activities* (PEES-LPA), was developed through expert review and numerous pilot procedures based on Bandura's Self Efficacy Theory (1977; 1982). The PEES-LPA was constructed and validated using the recommendations and guidelines presented in the previously mentioned literature, paired with Bandura's *Guide for Constructing Self-Efficacy Scales* (2006), and DeVillis's *Scale Development: Theory and Practice* (2012).

Following the establishment of content and construct validity through a focus group and expert review, 182 in-service secondary physical educators from the United States and Canada completed the resulting instrument, a 63-item survey. Exploratory Factor Analysis revealed a six factor model that accounted for 67.8% of the total observed score variance. Additionally, results revealed: (a) resulting factors showing simple structure that aligns with literature supporting the classification of lifetime physical activities (AAHPERD, 2013), (b) factors were composed of items that logically relate, and (c) internal consistency showed to be very high. In conclusion, the PEES-LPA appears to be an appropriate instrument for measuring self-efficacy perceptions of physical educators, though further revisions should be explored to help remove redundant items that may influence multicollinearity.

DEDICATION

To my wife, Jennifer

I dedicate this dissertation to my best friend, and outrageously loving and supportive wife, Jennifer. Thank you for your encouragement, and sacrifice while I pursued this dream. Your belief in me has made me a stronger person. I am a very lucky man.

To my Mom and Dad

I love and appreciate you both very much. You are both strong, hardworking, and dedicated people. You have taught me what it means to be a teacher, leader, husband, and father. You have never stopped supporting and believing in me. I cannot thank you enough.

ACKNOWLEDGEMENTS

I would like to first acknowledge my advisor, mentor, and friend, Dr. B. Ann Boyce. Thank you for you your constant input, guidance, support, and for allowing me to grow as a professional. Your passion for the field of physical education and teacher education is unrivaled, and I am blessed to have had the opportunity to learn from you. I would also like to thank Dr. Luke Kelly, a mentor and friend, for being my continual sounding board, and providing support both professionally and personally. I will miss our office conversations. I would also like to thank Dr. Martin Block for providing me feedback that challenged and shaped who I am as a professional; your work and commitment to adapted physical education inspires me to be a better professional. I also appreciate Dr. Ruth Ferree for challenging me to think outside of my physical education mindset, and explore the larger education landscape. I would like to thank Michael Hull and Dr. Erin Ottmar who aided tremendously with my construct development and data analysis.

Thank you to my wife, Jennifer, for being the most loving and supportive person. You have kept me strong through this journey, and I look forward to the next chapter in our lives. I am also grateful for my parents, for your unwavering love and support, and encouraging me to pursue my passions.

In addition, I want to thank my colleagues and friends: Teasha Jackson, EunHye Kwon, Jason Bishop, David VanBenschoten, Sean Healy, Andy Colombo-Dougovito, Joann Judge, Amanda Campbell, Andrea Taliaferro, and Jennifer Krause for laughs, feedback, and support through this UVA journey.

TABLE OF CONTENTS

List of Tables

List of Figures

List of Appendices

CHAPTER I

INTRODUCTION

Over the past few decades, a significant body of literature has accumulated to support the value regular physical activity has across one's lifespan (Corbin, 2002; U.S. Dept. of Health & Human Services-USDHHS, 2008). Unequivocal evidence has been found to support the direct relationship between regular physical activity and the reduction of chronic illness, as well as enhancement of personal wellness (Corbin, 2002, Lubans & Morgan, 2008; USDHHS, 2008). In addition, regular physical activity has resulted in a myriad of positive outcomes specific to children and adolescents, including increased physical fitness, mental health, self-esteem, self-concept, and the reduced risk of numerous lifestyle diseases (Hillsdon & Foster, 2003; Strong et al., 2005; USDHHS, 2008). Due to the significant benefits an active lifestyle can have on children and adolescents, the promotion of physical activity during these years is not only imperative, but has been identified as a global health priority (Lubans & Morgan, 2008; Pushka, Benaziza, & Porter, 2003).

Although there is a body of research to support the benefits of physical activity, unfortunately, evidence shows a decrease in activity levels as children progress in to adolescence and adulthood (American College of Sports Medicine-ACSM, 201; USDHHS, 2008). An accurate representation of this trend was exhibited in two nationally recognized studies measuring participant physical activity levels using accelerometers, a tool that records human movement. Results from these studies revealed: (a) by the age group of 6-11, less than half of both boys and girls were meeting daily physical activity standards, (b) substantially fewer adolescents met the 60 minutes-a-day standard compared to younger children ages 6-11 (See Figure 1), and that (c) girls crossed *below* the recommended physical activity standard at an earlier average year in comparison to boys (Nader, Bradley, Houts, McRitchie, & O'Brien, 2008; Troiano et al., 2008;).

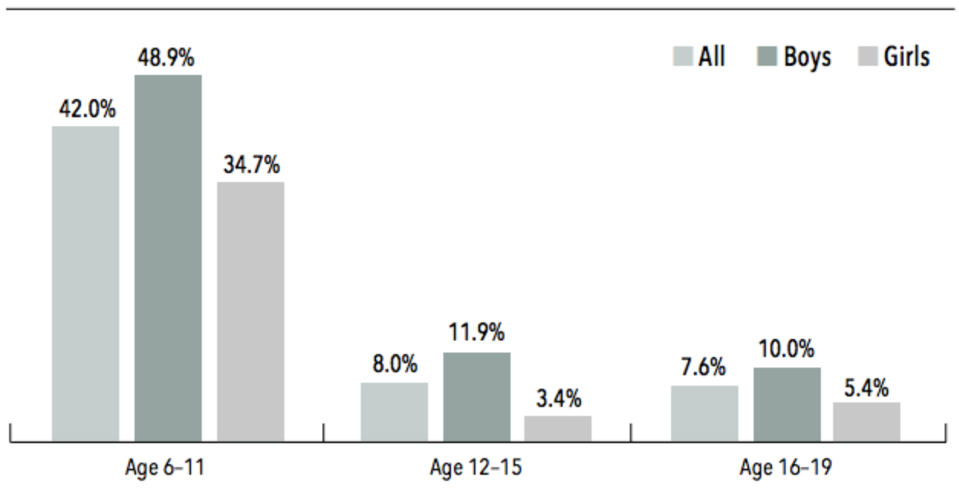


Figure 1. Children and adolescents (ages 6-19) getting at least 60 minutes per day of physical activity in the United States, 2003-2004. Adapted from "Physical activity in the United States measured by accelerometer," by R. P. Troiano, D. Berrigan, K. W. Dodd, L. C. Mâsse, T. Tilert, & M. McDowell, 2008, *Medicine & Science in Sport & Exercise, 40* (1), 181-188.

What makes the decrease of physical activity post childhood even more alarming is the well-documented relationship between physical inactivity, obesity, and chronic disease (Ross, Francis, BeLue, & Viruell-Fuentes, 2012). Individuals in today's fast-paced and technology-driven society are living more sedentary lifestyles by continually seeking less demanding ways to accomplish tasks (Amusa, Toriola, & Goon, 2012). Furthermore, there is a growing body of evidence to support that pairing the effects of physical inactivity with a sedentary lifestyle can have drastic effects on the deterioration of many bodily functions (Amusa, et al., 2012; Australian Institute of Health and Welfare, 2004; Pollock & Wilmore, 1990).

The prevalence of childhood obesity and physical inactivity has driven national organizations to lead federal initiatives to educate children and adolescents about making physical activity a lifelong commitment. Programs such as the First Lady's *Let's Move*, the American Alliance for Health, Physical Education, Recreation, and Dance's (AAHPERD) *Let's Move in School* initiative, the USDHHS's *Healthy People 2020*, and the *National Physical Activity Plan* all revolve around the promotion of physical activities that can be participated throughout one's entire life, or *lifetime physical activities*. A lifetime physical activity has been defined as any physical activity that may be readily carried over into adulthood due to the nature of being accomplished by one or two people, and that requires little structure, organization, and equipment (AAHPERD, 2013; Caspersen, Powell, & Christenson, 1985; Fairclough, Stratton, & Baldwin, 2002; Ross, Dotson, Gilbert, & Katz, 1985).

One common denominator that *all* of the above mentioned national initiatives addressed was school-based physical education as a vital component to the promotion of lifetime physical activity. School physical education has been found to be an vital medium for the promotion of lifetime physical activities for its ability to (a) offer classroom opportunities where students can regularly be physically active, and (b) creating experiences where students can participate in physical activities that can transfer outside of the school setting when transitioning into adulthood (Martin & Kulinna, 2003; Sallis, Simons-Morton, Stone, & Corbin, 1992).

To strengthen the link between physical education and lifetime physical activity promotion, the Center for Disease Control (CDC) brought together 17 leading field experts and 40 national agencies to develop the *Guidelines for School and Community Programs to Promote Lifelong Physical Activity in Young People* (CDC, 1997). This document summarized the literature relevant to promotion physical activity in school-based programs, and recommended guidelines for schools and community programs to encourage lifetime physical activity. Some illustrations of their literature-supported guidelines were:

- Planned and sequential physical education curricula should emphasize knowledge about the benefits of physical activity and the recommended amounts and types of physical activity needed to promote health (CDC, 1997, p. 13).
- Physical Education should help students develop the attitudes, motor skills, behavioral skills, and confidence they need to engage in lifelong physical activity (CDC, 1997, p. 13).
- Physical education should emphasize skills for lifetime physical activities rather than those for competitive sports (CDC, 1997, p. 13).
- Increase the proportion of physical education classes so students spend greater than 50% of the time being physically active, preferably engaged in lifetime physical activities. (CDC, 1997, p. 10).

Though individually each of these guidelines can be viewed as idealistic without administrative or legislative mandates to support them, clearly there is a shift in the language being used, and *lifetime physical activity* has risen to the top of dialogue, communication, and peer reviewed publications in physical education (AAHPERD, 2013). As a result of the emphasis being placed on schools for the promotion of lifetime physical activities, it then becomes essential to examine how confident physical educators are to develop and implement curriculum highlighting lifetime physical activities.

Research has shown that physical educators were major determinants not only in the activities in which students engage in during class, but students' achievement orientation and their overall increased perception of competence and enjoyment within physical activities (Martin & Kulinna, 2003; Ntoumanis & Biddle, 1999; Treasure & Roberts, 2001; Wallhead & Buckworth, 2004). Additionally, there is growing evidence to support that students who perceive themselves to be more competent in various physical activities tend to have an increased rate of participation in out-of-school physical activities in comparison to those who perceived themselves to be less competent (Carroll & Loumidis, 2001; Wallhead & Buckworth, 2004). This evidence affirms that the choices teachers make in relation to curriculum, lessons, content selection, and assessments influence physical activity participation (Gencay, 2009; Martin & Kulinna, 2003).

Knowing teachers play a large role in influencing student choices, the logical next step is to examine how physical educators promote lifetime physical activities. Researchers have recognized this by examining teacher confidence levels towards their ability to instruct specific activities, as well as the determinants that influence the choices they make, play an influential role in the application of physical activity instruction (Kujala, Kaprio, Sarna, & Koskenvu, 1998; Ross, et al., 1985; Sallis et al., 1992). Perceived confidence levels teachers have towards specific tasks are known as self-efficacy perceptions (Bandura, 1997).

Self-Efficacy Theory

Self-efficacy is defined as a set of beliefs, or expectations, about how competent a person feels in their ability to perform a particular task with a desired outcome (Bandura, 1997; 2006). One's self-efficacy beliefs can determine how they act, think, feel, and behave within specific situations (Dellinger, Bobbett, Olivier & Ellett, 2008). A major distinction grounded in this theory is that self-efficacy is not a measurement of one's *capacity* to perform a particular skill, but a measurement of one's *perception* of capability towards performing a particular skill (Bandura, 1997).

Therefore, an individual's perceived level of self-efficacy can be a significant determining factor in the performance of a particular skill (Pajares, 2002). For example, if an individual has the perception that they are highly capable at performing a particular task, their perception will then shape how they apply their knowledge and ability towards that task. Additionally, self-efficacy perceptions play a key role in human functioning because they not only affect direct behaviors, but it impacts other factors, such as goals, aspirations, outcome expectations, and perceptions of impediments in the social environment (Bandura, 2006). With this reasoning, an individual's behaviors can often be better predicted by the beliefs that they hold about their capabilities, rather than by what they are actually capable of accomplishing (Pajares, 2002). Efficacy perceptions can then act as determinants for the individual's application of the specific knowledge and skills.

Due to the task- and situational-specific defining characteristics of the self-efficacy construct, Bandura (1997; 2006) has affirmed that there can be no one all-purpose measure for perceived self-efficacy. Self-Efficacy measurement must be specifically aligned to activity domains, and assess multidimensional ways in which self-efficacy beliefs operate within the selected activity, thus, linking factors that demine quality of functioning in the domain (Bandura, 2006).

Although general teacher self-efficacy has been explored at great length (Bandura, 1997), unfortunately few researchers have examined teacher self-efficacy among physical educators, or efficacy specific to teaching physical education (Humphries, Hebert, Daigle, & Martin, 2012; Martin & Kulinna, 2003). This is especially concerning for two reasons: (a) the obesity epidemic has placed increased emphasis on physical educators to become the gateway for students developing physically active lifestyles (Pan, 2012), and (b) further demand on accountability and rethinking of curricular strategies in physical education (Edgington, Kirkpatrick, Schupbach, Phillips, & Chen, 2011). If future research is to continue to develop an understanding of physical education teacher self-efficacy, the assessment instruments used (or developed) should be specific to the content of physical education.

Martin and Kulinna (2003), while creating and validating their *Physical Education Teacher Physical Activity Self-Efficacy Instrument* (PETPAS), emphasized that if teachers have weak self-efficacy towards "teaching physically active lessons, it would suggest that policy makers' vision of physical education as a public health vehicle might not be shared by physical educators" (p. 268). The teachers who are educating children about health, wellness, and lifetime physical activity are assumed to be the appropriate facilitator for shifting student perceptions, attitudes, and actions. If teachers lack the self-efficacy beliefs to appropriately advance student development and to promote a quality physical education experience, then how can we depend on physical education as a main source for remedying the current health epidemic?

Measuring Self-Efficacy

Numerous issues have plagued teacher self-efficacy research, a majority of those issues stemming from inconsistent validity and reliability procedures conducted during the analysis of psychometric properties (Tschannen-Moran & Woolfolk-Hoy, 2001). Bandura (1997) found that the construct of self-efficacy has been inadequately assessed with one-item-scales that have failed to demonstrate agreement between the measurement scale and the specific behavior being targeted. Bandura also went on to address the fact there is a definitive need for sound measurements in education that are strongly linked to the theoretical groundwork of Self-Efficacy Theory, and are follow a prescriptive path of scale development (Bandura, 1997).

In regards to context driving the reliability across scale development, DeVellis (2012) remarked that scales are "reliable to the extent that they consist of reliable items that share a common latent (construct) variable" (p. 47). The evaluation of scale reliability and validity is known as psychometric testing.

Psychometric testing is used in various fields, including education and psychology research, to define properties of the items that are in included in a survey (DeVillis, 2012). Survey instruments should undergo a full range of psychometric testing before any decisions are made based on their ability to measure a particular construct. Because no two scales or psychometric evaluation procedures are the same, there is a great amount of variability across approaches researchers have taken towards determining validity and reliability of survey instruments.

Of the handful of peer-reviewed self-efficacy instruments that are present in the literature specific to physical education teachers, replication issues may arise if pursued by future researchers. These limitations stem from: (a) instruments being too narrowly focused to reach global appeal, (b) survey items being written with global context which lacks specificity towards individual teaching behaviors, thus not controlling for external variables (e.g. general teaching efficacy in physical education, efficacy to include students with special needs in physical education, efficacy to integrate technology in physical education, etc.), (c) the use of double-barreled questions, and (d) lack of adequate psychometric analysis.

There currently is no universal consensus on required psychometric properties when reporting reliability and validity procedures during instrument construction and evaluation (American Education Research Association, 1999). This has lead to inconsistencies on the defining characteristics for reporting psychometric reliability and validity evidence during scale development. It is unfortunately very common for peer-reviewed authors to describe the "reliability of the test", or stating, "the test is reliable." Such statements contribute to the confusion and misunderstanding of the concept and features of score reliability and validity (Vacha-Haase, 1998). The discrepancies in reporting psychometric information has failed to provide a unified opinion on what is deemed to be an "appropriate or acceptable" amount of psychometric testing. What has resulted is a tendency towards inclusion of an abundance of psychometric testing, which often leads to higher evidence for reliability and validity (American Education Research Association, 1999).

Due to the emphasis the physical education field has recently placed on instruction of lifetime physical activities, there needs to be a way measure how confident teachers are in administering instruction in these activities. Presently there is no measure that would address such self-efficacy perceptions, primarily because past instruments are: (a) not specific to the construct of lifetime physical activities, (b) too global in nature, or (c) are without strong enough evidence for validity and reliability. In order to address this need, a psychometrically sound self-efficacy scale must be created addressing teacher perceptions towards instruction of lifetime physical activities.

Statement of the Problem

Self-efficacy perceptions towards instruction in physical education is still largely unexplored when it comes to the dynamic nature of human motivation and behavior towards the promotion of physically active lifestyles. Both Martin and Kulinna (2003) and Humphries et al. (2012), have emphasized a need for further examination of self-efficacy beliefs of physical educators, as well as the need for future research to address how self-efficacy perceptions can be a strong predictor of how teacher preparation and specific experiences influence instructional decisions.

Due to the emphasis numerous national, state, and local organizations have placed on the promotion of lifetime physical activity in our school physical education programs, it is imperative that researchers start to examine how this implementation is taking place. As a result, a valid measure is needed to assess how teachers perceive their abilities to teach lifetime physical activities. Currently, there is no existing valid instrument that measures self-efficacy perceptions of physical educators towards the instruction of lifetime physical activities. To address the gap in the research, as well as the shortcomings of previously constructed physical educator self-efficacy instruments, a new scale with strong evidence for validity and reliability is needed.

Purpose of the Study

This study was conducted in response to a need for a psychometrically sound instrument measuring physical educator self-efficacy perceptions towards instruction of lifetime physical activities. The purpose of this study was to begin to determine the validity and reliability of a newly constructed instrument. This study involved the preliminary procedures of developing an instrument, followed by the validation of the instrument and an exploration of the data for evidence to support validity.

Overview of Research Questions

Two research questions (RQ) were addressed in the current study:

RQ1: Does the Physical Educator Efficacy Scale for Teaching of Lifetime Physical Activities (PEES-LPA) provide evidence for validity in measuring physical educator self-efficacy perceptions towards instruction of lifetime physical activities?

RQ2: Does the Physical Educator Efficacy Scale for Teaching of Lifetime Physical Activities (PEES-LPA) provide evidence for reliability in measuring physical educator self-efficacy perceptions towards instruction of lifetime physical activities?

Definition of Terms

Cronbach's Alpha. Cronbach's alpha is a measure of internal consistency, that is, how closely related a set of variables are as a group (DeVellis, 2012).

Eigenvalue. Eigenvalues are each of a set of values of a parameter for which a differential equation has a nonzero solution (an eigenfunction) under given conditions (Field, 2009).

Exploratory Factor Analysis. A statistical method used to uncover the underlying structure of a relatively large set of variables (Field, 2009).

Factor Loadings. It is term used to refer to factor pattern coefficients or structure coefficients (Field, 2009).

Internal Consistency. Measures whether several items that propose to measure the same general construct produce similar scores (DeVellis, 2012).

Latent Variable. A latent variable is the un-observable construct that a social science researcher is trying to measure (Field, 2009).

Mastery Experience. The most powerful source of self-efficacy; it is one's interpretations of his or her own previous, authentic experiences performing a particular task (Bandura, 1997).

Physical Education Teacher Education (PETE). A higher education kinesiology program of study, it prepares students to become physical education teachers and leads to teacher certification.

Psychological States. The fourth most effective way to develop self-efficacy; it is the perception and interpretation of emotional and physical reactions, such as stress, anxiety, fear, etc. (Bandura, 1997).

Principal Axis Factoring. A method of factor analysis in which the factors are based on a reduced correlation matrix using a priori communality estimates. That is, communalities are inserted in the diagonal of the correlation matrix, and the extracted factors are based only on the common variance, with specific and error variances excluded (Field, 2009).

Reliability. The ability of a measure to produce consistent results when the same entities are measured under different conditions (Field, 2009).

Scree Plot. In factor analysis, a graph of all eigenvalues (y-axis) compared to the factor they are associated with (x-axis)(Field, 2009).

Self-Efficacy. "Beliefs in one's capabilities to organize and execute the courses of action required to produce given attainments" (Bandura, 1997, p. 3).

Self-Efficacy Theory. A theory that proposes self-efficacy as a common cognitive mechanism for mediating people's motivation and behavior (Bandura, 1997).

Social Persuasion. The third most effective way to develop self-efficacy, where verbal persuasion by credible, trustworthy, or expert persons is given (Bandura, 1997)

Sources of Self-Efficacy. The hypothesis that states that self-efficacy beliefs develop, can be instilled, and can be strengthened as people interpret information from four sources: mastery experience, vicarious experience, social persuasion, and psychological states (Bandura, 1997).

Validity. Considers whether a test measures what it is supposed to measure, and refers to the appropriateness and usefulness of specific inferences made from test scores (Thorndike, 1997, 2005; DeVellis, 1991).

Vicarious Experience. The second most powerful source of self-efficacy, where one estimates his or her own capabilities in relation to the performance of others (Bandura, 1997); provided in the form of feedback for this study.

Delimitations

The study is delimited in the following areas:

1. Only current in-service physical education teachers will be included in the study.
2. Physical educators' self-efficacy beliefs will be explored in regard to instruction of lifetime physical activities.
3. Only physical education teachers who volunteer for the study and complete the survey will be included in the study.

Limitations

A limitation of a study concerns the generalizability of its findings. The following are limitations of the study:

1. Behaviors are self-reported, and therefore may not be an accurate

measurement of actual behavior.

2. Participants who volunteer for the study may not be a representative sample of the population, as individuals with higher levels of self-efficacy may be more willing to complete the survey.
3. Some participants will be invited to participate in the study by use of a listserve connected to membership in a professional organization or through county physical education directors. The results might be biased by this selection factor, as the teachers who choose to participate in these organizations and activities may be more motivated to maintain and improve their teaching skills and more up-to-date with current issues in the field, and therefore may not necessarily be representative of the population.
4. Since the beliefs of physical education teachers are being explored, it cannot be assumed that findings will generalize to other content areas.
5. Since physical educators' self-efficacy beliefs are being explored in regard to instruction of lifetime physical activities, it cannot be assumed that findings will generalize to overall instruction in physical education or to specific activities.
6. Participants for this study are in-service physical education teachers. Findings will not necessarily generalize to pre-service teachers.
7. Physical education teachers may respond in ways they feel are more socially acceptable rather than indicating their true beliefs.
8. Due to inherent differences in teaching practices at various schools, there may be variation in the extent to which inclusion is implemented in physical

education.

Statement of Significance

The results of this study will provide an understanding of the relevance of self-efficacy theory in regard to teaching lifetime physical activities in physical education by developing and starting the validation and reliability process of a proposed survey instrument. Due to the need placed on instruction of lifetime physical activities in physical education, the PEES-LPA instrument will add to the existing instruments in the field and will be the only instrument currently available to specifically explore self-efficacy beliefs of in-service physical educators toward instruction of lifetime physical activities. Finally, future results could help identify methods to improve self-efficacy and positive teaching behaviors of physical educators toward instruction of lifetime physical activities and could potentially guide future physical education teacher education and professional development programming.

CHAPTER 2

REVIEW OF LITERATURE

The primary purpose of this study was to create and begin to validate a self-efficacy instrument measuring physical educator perceptions towards instruction of lifetime physical activities. This chapter provides a review of the relevant literature and is divided into four major sections: (a) lifetime physical activity (b) self-efficacy theory, (c) methodology for scale development, and (d) measurement of physical education teacher self-efficacy. Each topic will be examined and pertinent literature will be summarized, with the chapter conclusion emphasizing importance of self-efficacy research, and offering support for the development of a new efficacy instrument measuring physical education teacher efficacy perceptions.

Section I: Lifetime Physical Activity

The first section of this literature review will focus on the examination of lifetime physical activity and its context in physical education. This section will be presented in the following segments: (a) value of lifetime physical activity, (b) lifetime physical activity in physical education, (c) defining lifetime physical activities, (d) conclusion of lifetime physical activity.

Value of Lifetime Physical Activity

The rise in prevalence of overweight and obese individuals, particularly with children and adolescents, is one of the most disturbing health issues facing

Americans today. In the United States, approximately 35 percent of adults, and 17 percent of children and adolescents are affected by obesity (Bell, Rogers, Dietz, Ogden, & Shuler, 2011). What makes these statistics even more alarming is that in the past 30 years, childhood obesity has doubled in children, and tripled in adolescents (Ogden, Carroll, Kit, & Flegal, 2012).

One major cause for the rise in obesity is lack of physical activity (USDHHS, 2011). Physical activity is defined as bodily movement produced by the skeletal system resulting in energy expenditure (Caspersen, et al., 1985). On its basic level, physical activity can involve such things as walking and jogging, doing household chores, dancing, or yard work. Though physical activity opportunities are readily available and easy to access, many physiological, behavioral, and psychological variables affect physical inactivity. Sallis, et al., (1992) deduced from their research that there are 10 major reasons people cite for not adapting a more physically active lifetime: (a) not enough time, (b) inconvenient, (c) lack of self-motivation, (d) not enjoyable, (e) boring, (f) low self-efficacy, (g) fear of injury, (h) lack of self-management skills, (i) lack of encouragement/support, (j) lack of facilities near home/office. Bell et al. (2011), in their examination of childhood obesity in the U.S., concluded that to shift the negative connotations associated with physical activity, intervention programs should focus on environmental determinants (e.g., school, home, church) altering both food choices and physical activity opportunities.

In 2010, the Center for Disease Control (CDC) conducted a national examination of physical activity levels among Americans based on recommended physical activity guidelines (USDHHS, 2008). In the published report, the *State*

Indicator Report on Physical Activity, 2010, the CDC emphasized, overall, Americans are not getting enough physical activity. More specifically they highlighted:

- Less than half (48%) of adults meet the physical activity guidelines.
- Less than 30% high school students get at least 60 minutes of physical activity every day.
- The Southern U.S. states are more likely to be inactive, compared to the West, Northeast, and Midwest regions of the country.
- Younger adults are more likely to meet physical activity requirements then older adults.
- Adults with more education are more likely to meet physical activity guidelines for aerobic activity than adults with less education. (CDC, 2010).

To further illustrate the trend towards physical inactivity, two different studies published in 2008 used accelerometer data from national studies to examine physical activity levels among children and adolescents (Nader, et al., 2008; Troiano et al., 2008). An accelerometer is a movement monitor that has the ability to capture intensity of physical activity, and distinguishes itself from pedometers by its ability to classify differences in movement patterns (i.e., walking vs. running), and separate out movement outside of normal human functioning (e.g., vibration from a car) (Troiano, et al., 2008). Though studies examined accelerometer data from separate national data sets (*2003-2004 National Health and Nutrition Examination Survey*; *1991-2007 National Institute of Child Health and Human Development Study of Early Child Care and Youth Development*) similar results showed that:

- Males are more active than females.
- Physical activity declines dramatically across age groups between childhood and adolescence.
- Girls crossed below the recommended physical activity standards at an earlier average year compared to boys.
- There is a huge drop-off in physical activity after the age of 11.

(Nader et al., 2008; Troiano et al., 2008)

In 2012, the CDC published the 2011 National Youth Risk Behavior Survey (YRBS), which collected self-reported data from high school students across the U.S. based on physical activity levels. Coinciding with the data found by Troiano, et al. (2008) and Nader et al. (2008), the CDC found that many youth are not meeting the recommended 60 minutes of daily physical activity. Their results also showed that 29 percent of high school students participated in at least 60 mutes of physical activity, and that 14 percent of students did not participate in any kind of physical activity at all.

What the literature clearly demonstrates is that as people progress towards adolescence and adulthood physical activity levels drastically decrease. This tendency is alarming primarily due to the irrefutable evidence of the effects physical activity can have on numerous health benefits and chronic diseases (Ross, et al., 2012). Warburton, Whitney, and Bredin (2006) and Janssen and LeBlanc (2010) both conducted systematic analyses examining the empirical evidence of the health benefits of physical activity. Results from the two analyses showed that routine physical activity is linked to improvement of body composition, mental health, bone strength, and reduction in the risk of obesity, cardiovascular disease, and diabetes (Janssen & LeBlanc, 2010; Warburton et al., 2006). Furthermore, both analyses concluded that there is a direct linear relationship between volume of physical activity and overall health status (i.e., the most physically active people have the lowest risk of chronic disease) (Janssen & LeBlanc, 2010; Warburton et al., 2006). Much of the literature published by the CDC (2010; 2012) validates the findings and conclusions from these two systematic analyses.

Due to the glaring need for reformation on views Americans have towards physical activity, many government and national initiatives have been enacted to promote physical activity throughout a lifetime. Programs such as the First Lady's *Let's Move!*, the American Alliance for Health, Physical Education, Recreation, and Dance's (AAHPERD) *Let's Move in School* initiative, the USDHHS's *Healthy People 2020*, the National Institute of Aging's *Go4Life*, and the *National Physical Activity Plan* all revolve around the promotion of physical activities throughout one's lifetime. A summary of the physical activity initiatives can be viewed in Table 1.

Lifetime Physical Activity in Physical Education

Aside from the *Go4Life* physical activity program for older adults, *all* of the other initiatives address school-based physical education an instrumental player in the promotion of lifetime physical activity. Additionally, the American Academy of Pediatrics, NASPE, American Health Association, USDHHS, the President's Council on Physical Fitness and Sport, and the CDC have all gone on record supporting physical education as essential in the promotion of lifetime physical activity (LeMasurier & Corbin, 2006). According to the Snyder and Dillow (2012), more than 55 million

Table 1

National and Government Initiatives Addressing Physical Activity in the United States

Name	Goal/Purpose	Focuses
Let's Move!	Campaign to end childhood obesity	•Creating a healthy start for children •Empowering parents and caregivers •Providing healthy food in schools •Improving access to healthy, affordable foods •Increasing physical activity
Let's Move in Schools	Make school healthier places to learn by providing quality nutrition, integrating physical activity, and teaching about the importance of embracing a healthy active lifestyle	•Creating a school health advisory council •Join Healthier US Schools Challenge •Make School a healthy worksite •Incorporate nutrition education and physical education in to curriculum •Plant a school garden
Healthy People 2020	Improve health, fitness, and quality of life through daily physical activity.	•Physical activity in childcare settings •Decreasing television and computer usage •Focusing on the importance of recess and physical education in public and private elementary schools •Improve access to facilities that promote physical activity
National Physical Activity Plan	One day, all Americans will be physically active and they will live, work, and play in environments that facilitate regular physical activity.	•Provide access to and opportunities for high-quality, comprehensive physical activity programs •Develop and implement state and school district policies requiring school accountability for the quality and quantity of physical education and physical activity programs. •Linking youth with physical activity opportunities in schools and communities. •Ensure that early childhood education settings for children ages 0 to 5 years promote and facilitate physical activity. •Provide access to and opportunities for physical activity before and after school. •Encourage post-secondary institutions to provide access to physical activity opportunities, including physical activity courses, robust club and intramural programs, and adequate physical activity and recreation facilities.
Go4Life	Designed to help older adults fit exercise and physical activity into your daily life	•Motivating adults to become physically active for the first time •Motivating adults to return to exercise •Build more exercise and physical activity into weekly routines

children were expected to attend public and private school in the fall of 2012, with those students attending an average of 6 to 7 hours a day. This makes school-based physical education an ideal setting for the promotion of lifetime physical activity.

Additional evidence to support the link between physical education and lifetime physical activity promotion can be seen in the work of Gordon-Larsen, McMurray, and Popkin (2002). They examined longitudinal data addressing hours per week of inactivity and times per week of moderate-to-vigorous physical activity. Regression results from their study show a significant relationship between participation in physical education and physical activity patterns, as well as an indication that physical education classes may represent the single best opportunity for many adolescents to engage in regular physical activity.

Quality of instruction has also been found to be a controlling factor in the promotion of lifetime physical activity. In 2012, the USDHHS published the *Physical Activity Guidelines for Americans Midcourse Report: Strategies to Increase Physical Activity Among Youth*. Included in this report was a systematic review examining the implementation of physical activity promotion in physical education. The USDHHS's results indicated strong evidence supporting curricular and instructional decisions from qualified physical educators being a primary influential variable in the promotion of physically active lifestyles.

Currently, there are a number of national initiatives that have addressed secondary physical education curriculum as being the primary vehicle for the promotion of lifetime physical activity (AAHPERD, 2013; CDC, 1997; 2010; NASPE, 2009). In the United States, the national standards for K-12 physical education

(NASPE, 2009) specifically state, "the goal of physical education is to develop physical literate individuals who have the knowledge, skills, and confidence to enjoy a lifetime of healthful physical activity" (p. 5). Though different secondary physical education programs may approach teaching lifetime physical activities in various ways, it is clear from the NASPE (2009) guidelines that a lifetime of physical activity is the catalyst to prolonged wellness.

Therefore, the next logical step in addressing the effectiveness of the promotion of lifetime physical activity in physical education curriculum, is determining if qualified physical educators are confident in their ability to deliver such skills and competencies to their students (Carroll & Loumidis, 2001; LeMasurier & Corbin, 2006, Treasure & Roberts, 2001; USDHHS, 2008). Confident and qualified physical educators help engage and expose students to a wide range of physical activities that are enjoyable and build self-efficacy so they can continue to stay active throughout a lifetime (Sallis, et al., 1992; USDHHS, 2008).

Defining Lifetime Physical Activities

Researchers such as Fairclough, et al., (2002), have suggested that children who find physical activity a positive experience from an early age are much more likely to sustain physical activity in to adulthood. As a result, there has been a strong importance placed on physical education curriculum focusing on the promotion of physical activities that aligning with lifelong participation, opposed to team sport and striking and fielding games (Fairclough, et al., 2002; American College of Sports Medicine, 2011; AAHPERD, 2013). Compared to team or invasion games, lifetime physical activities are now seen as having much better carry-over value as students transition into adulthood (Fairclough, et al., 2002; Sallis, et al., 1992).

Ross, Dotson, Gilbert, & Katz (1985) define *lifetime physical activities* as any physical activity readily carried over into adulthood due to the nature of being accomplished by a minimum of one or two people, and require little structure, organization, and equipment. Though numerous activities appear to fit this definition, a recent publication by the American Alliance for Health, Physical Education, Recreation and Dance (AAHPERD, 2013) helped clarify what activities constituted being defined as a *lifetime physical activity*. The operational definition for this study pertaining to lifetime physical activities will align with AAHPERD's (2013) recommendation. AAHPERD (2013) state that lifetime physical activities are any activity that falls within the following seven categories:

1. Outdoor Pursuits (e.g., hiking, backpacking, snow/water skiing, mountain biking).
2. Fitness Activities (e.g., running, cycling/biking, yoga, weight/resistance training).
3. Dance & Rhythmic Activities (e.g., ballet, modern, line, social and square)
4. Aquatics (e.g., swimming, diving, water aerobics)
5. Individual Performance Activities (e.g., gymnastics, track and field, self-defense)
6. Net/Wall Games (e.g., tennis, pickleball, badminton)
7. Target Games (e.g., golf, archery, bowling)

Prior to the AAHPERD (2013) publication, there has been little consensus to the operational definition of lifetime physical activity. Confusion can arise from the use of closely related terminology, which as a result misconstrues the phrase's actual definition.   First, the use the terms *lifelong physical activity* and *lifetime physical activity* have been used interchangeably. Though these two phrases are synonymous of one another, the term *lifetime physical activity* is more widely accepted within the scholarly community for physical education, therefore it will be the delimited terminology used for this study (AAHPERD, 2013; NASPE, 2008).

The second clarification is the difference between *physical activity* versus *physical exercise*. These two phrases have often been incorrectly used interchangeably, primarily due to the fact both contain a number of common elements. Exercise, however, is not synonymous with physical activity, and is a subcategory of physical activity (Caspersen et al., 1985). Physical exercise is "physical activity that is planned, structure, repetitive, and purposive in the sense that improvements or maintenance of one or more components of physical fitness is an objective" (Caspersen et al, 1985, p. 128).   The term lifetime physical activity was selected over that of lifetime physical exercise due to exercises being more centered on fitness movements that are often repetitive in nature (e.g. strength training, interval training), which by definition excludes daily physical movements (e.g. gardening, biking to work) and physical activities, such as hiking, yoga, or track an field.

Conclusion for Lifetime Physical Activity.

It is obvious from the widespread literature and national initiatives that obesity and chronic health problems are becoming an epidemic in the United States. Physical inactivity at all age levels is one of the major reasons causing the increase in heath concerns. School-based physical education has been cited as an essential place for students to gain positive experiences and exposure to physical activities that can be carried throughout a lifetime. Though physical education is an excellent outlet for students to participate in a large array of lifetime physical activities, the operational definition of *lifetime physical activity* is still unclear in many physical education circles. With the recent AAHPERD push towards physical education curriculum emphasizing lifetime physical activities, the operational definition of what constitute appropriate physical activities for lifelong participation is becoming clearer.

Section II: Self-Efficacy Theory

The purpose of this second section was to review relevant literature on self-efficacy, the theoretical framework for this study. This chapter provides a review of the relevant literature and is divided into five major sections: (a) theoretical perspective (b) self-efficacy and human functioning, (c) self-efficacy and psychological processes, (d) sources of self-efficacy, and (e) measuring self-efficacy perceptions.

Theoretical Perspective

Self-efficacy theory served as the theoretical framework for the investigation of physical educator's perceptions towards the instruction of lifetime physical activities. Bandura (1986; 1997) has stated that there are two major resources that a person must have to be successful in performing a task: skill/knowledge and self-efficacy. In an academic context, personal self-efficacy perceptions of teachers have been found to strongly influence instructional decisions, student achievement, classroom decisions, as well as their orientation towards the educational process (Bandura, 1982). From this viewpoint, this researcher hypothesizes that increasing the self-efficacy of physical education teachers towards instruction of lifetime physical activities actually could be a valuable method in helping students gain exposure and self-efficacy towards physically active lifestyles.

Human competencies can manifest and be developed in many different ways. One cannot simply master everything; this is far too time consuming, and would require a vast dedication of resources and effort (Bandura, 1997). Thus, individuals cultivate and develop their competencies based on chosen pursuits. "The particular patterns of competencies they acquire are products of natural endowment, sociocultural experiences, and fortuitous circumstances that alter the course of developmental trajectories" (Bandura, 1997, p. 36). Self-efficacy theory embraces the diversity in human capabilities, and supports that the efficacy belief system not as a collective trait, but as a differentiated set of self-beliefs linked to specific domains of functioning (Bandura, 1997).

Albert Bandura, the theory's originator, has defined self-efficacy as "beliefs in one's capabilities to organize and execute the courses of action required to produce given attainments" (1997, p. 3). In other words, self-efficacy is an individual's judgment of their ability to successfully perform a skill/action within a specific context (Pajares, 2002). Self-efficacy theory is one of the major elements of Bandura's Social Cognitive Theory (1986), which describes the interactions between personal factors, behavior, and environmental influences in the context of human functioning (Pajares, 2002). A major distinction in self-efficacy theory is that there are definitive differences between possessing skills, and being able to integrate them into an appropriate course of action during difficult situations (Bandura, 1993). For example, people can fail to perform optimally in a particular situation even though they know what to do and possessed the required skills to do so (Schwartz & Gottman, 1976). "Self-referent thought activates cognitive, motivational, and affective processes that govern the translation of knowledge and abilities in to proficient action" (Bandura, 1997, p. 37). As a result, perceived self-efficacy does not focus on the skills you have, but how you believe you can apply those skills in diverse situations (Bandura, 1993; 1997).

Bandura (1977; 1982; 1997) found an individual's perception of self-efficacy to be multi-dimensional and consisting of two main components: (a) efficacy expectations, and (b) outcome expectations. Efficacy expectation is the belief that one has the ability to successfully execute the behavior required to achieve a specific outcome, while outcome expectancies are an individual's estimate that a give behavior will lead to certain outcomes (Bandura, 1977; 1997). Although Bandura originally hypothesized that self-efficacy and outcome expectancies were both major contributors to motivation, he later suggested that self-efficacy would have larger influence due to "the types of outcomes people anticipate depend largely on their judgments of how they will be able to perform in a given situation" (Bandura, 1986, p. 392). He further acknowledges the need to differentiate the two because one can believe that they a particular course of action will produce desired outcomes, but if they have self-doubt regarding their capability to perform the necessary skill/behavior that individual may choose to avoid any attempt altogether (Bandura, 1977; 1986).

Self-efficacy perceptions have been found to vary on three different performance implications: magnitude, generality, and strength. The magnitude of self-efficacy refers to its dependence on the difficulty of a specific task, or the level of task difficulty a person believes he/she can attain. Generality pertains to the transferability of self-efficacy perceptions across various situations, such as from soccer to basketball. Strength of perceived self-efficacy, the basis of most self-efficacy instruments, measures the amount of one's certainty about a performance in a particular task (Bandura, 1977; 1982).

Self-Efficacy and Human Functioning

An individual's efficacy beliefs hold a lot of power over their human agency, not only directly, but also by impact on other determinants such as goals and aspirations, outcome expectations, affective proclivities, and perception of impediments and opportunities in the social environment (Bandura, 1982; 1997; Pajares, 2002). Unless people feel that their actions can produce positive or influential results based on the outcomes they desire, they will have very little incentive to attempt or pursue action in facing those difficulties (Pajares, 2002; Bandura 1997). In general, individuals often do not undertake tasks that they feel are outside of their abilities (Caprara, Barbaranelli, Pettita, & Rubinacci, 2003). If one feels they are capable of mastering a specific task, it then leads the individual to see the constraints of the task as challenges, rather then impossible obstacles (Bandura, 1993; Caprara et al., 2003). "These beliefs prevent preoccupations and ruminations from interfering with carrying out plans and instead help one focus on the problems and best use one's capabilities and the available resources" (Caprara et al., 2003, p. 16).

People who have high levels of self-efficacy are more inclined to approach difficult tasks as challenges to be mastered, foster interest, set challenging goals and stay committed to those goals, have higher effort, and quickly recover after failure (Bandura, 1997; 1977; 1993). Efficacious outlooks also act as an agent in producing higher probability for personal accomplishment, and reducing stress, anxiety, and vulnerability to depression (Bandura, 1993). Conversely, inefficacious people shy away from difficult tasks, are easily discouraged by impediments, and tend to believe tasks are more difficulty then actually are, have lower aspirations, dwell on deficiencies, and focus on adverse consequences of failure (Bandura 1997; Pajares, 2002). Stajkovic and Luthans (1998) conducted a large meta-analysis (N= 114) to examine the relationship between self-efficacy and human functioning associated with work related performance. Their report showed that that self-efficacy explained a 28% increase in work-related performance as compared to organizational behavior modification (17%), feedback intervention (13.6%), and goal setting (10.9%) in a two-decade meta-analysis.

In relation to human functioning of teachers, self-efficacy has also shown to influence behaviors such as: increase use of various teaching modalities (Allinder, 1994, Guskey, 1988; Stein & Wang, 1988), persist longer with students who are struggling (Gibson & Dembo, 1984), are less critical of student errors (Ashton & Webb, 1986; Gibson & Dembo, 1984), greater classroom-based decision making (Hoy & Woolfolk, 1993; Ward, 2005), and an overall greater enthusiasm towards teaching (Allinder, 1994; Hall, Burley, Villeme, & Brockmeier, 1992). These trends to help support the need for continued work in teacher self-efficacy research, but what they fail to address, is the specificity of teaching physical education. For example, the Gibson and Dembo (1984) *Teacher Efficacy Scale* (TES), arguably the most widely used self-efficacy scale over the past 30 years, presents broad items that lack in ability to address task constraints specified by Bandura (1977; 1982; 1997). Two example items from TES asks the following: (a) "when the grades of my students improve it is usually because I found more effective teaching approaches" (p. 573), and (b) "when a student does better than usual, many times it is because I exerted a little extra effort" (p. 573). Though these questions do address general teacher efficacy, the primary focus of TES, they lack the specificity necessary to measure physical educators.

Self-Efficacy and Psychological Processes

Bandura (1993; 1997) indicates that there are four psychological processes in which self-efficacy beliefs produce their results: cognitive, motivational, affective, and selection. In relation to cognitive processes, self-efficacy beliefs affect thought patterns that can enhance or inhibit future performance. As a result, self-efficacy perceptions influence personal goals setting and self-appraisal of capabilities. For example, a physical education teacher with a higher perceived self-efficacy will set high teaching goals and standards, challenge personal growth, and remain committed to those goals and personal challenges (Bandura, 1993). Motivational processes are impacted by self-efficacy by resilience to failure, perseverance, and persistence to attain selected goals (Bandrua, 1997). Self-efficacy influences the affective processes by influencing how much stress and depression one experiences in a difficult situation, as well as their ability to exercise control over stressors. Finally, self-efficacy influences selection processes. This result is a summation of cognitive, motivational, and affective processes and the influence self-efficacy has on the activities people choose to partake. For example, a physical educator who has no educational or participatory experience with aquatics, thus having low self-efficacy towards the instruction of aquatics, may refrain from choosing to include swimming in a physical education curriculum (Bandura, 1993; 1997).

Sources of Self-Efficacy

Bandura (1977; 1997) outlined that self-efficacy beliefs develop in the response to four sources of information: enactive mastery experiences, vicarious experiences, verbal persuasion, and psychological and affective states. Information gathered from these four sources combined with personal, social, and situational factors to influence how an individual reflects up on and interprets their experiences (Bandura, 1977, 1982).

Mastery experiences. Mastery experiences, also known as enactive mastery experiences, have been proposed to be the most influential source of information in the development of self-efficacy perceptions (Bandura, 1977; 1997). Mastery experiences have been defined as one's interpretations of his or her own previous authentic experiences to perform a particular task. Self-efficacy towards a task is also highly dependent on success. Successful experiences in performing an action build robust perceptions of self-efficacy, whereas failures work to undermine self-efficacy (Bandura, 1997).

Positive mastery experiences can also work as a reinforcement of previous self-efficacy beliefs. For example, if a physical educator was grew up having success playing the sport of tennis, attended a teacher education program focused on development of skill themes, movement patterns, and pedagogical content around racquet sports, one could hypothesize that this teacher would have a high self-efficacy perception towards the instruction of tennis. These numerous first-hand experiences helped shape and build upon self-efficacy perceptions. The only stipulation to influence mastery experiences have on self-efficacy is that the individual must reflect and process their abilities and experiences as they happen. Without time to process over an experience, or if the experience was too easy to register as a task constraint, there is not cognitive processing required (Bandura, 1994; 1997).

Vicarious experiences. Vicarious experiences, also known as modeling, occur when an individual views others, or social models, performing a particular task with success (Bandura, 1994). The individual then assesses his or her own abilities in comparison to the performance of the social model. Vicarious experiences are often most successful when the individual views the social model as similar in ability or skill (Bandura, 1993; 1997). If an individual feels that he or she possesses similar ability levels to the social model, then success had by the social model acts as an enabler for the individual to perform the same task with a similar outcome (Bandura, 1993; 1997). Vicarious experiences convey the same basic concept as mastery experiences; however, the individual carries out the task in mastery experiences, and the individual experiences the task through the eyes of a second party in vicarious experiences.

Social persuasion*.* The third source of self-efficacy information is known as social (or verbal) persuasion. Social persuasion occurs when an individual is given specific feedback by others regarding his or her ability to perform a specific task (Bandura, 1993; 1997). Bandura (1997) argued that individuals who received quality encouragement, are more likely to exert greater effort to become successful then if they filled with self-doubt. Supportive messages can aid in boosting one's self-confidence and effort, especially when paired with instruction and conditions that foster success (Bandura, 1997). For example, if a pre-service teacher was given encouragement about their demonstration and verbal description a particular weight training skills, the encouragement can lead to improvement over the teachers' self efficacy to do so.

Bandura (1994; 1997) states that there are two major cautions when addressing social persuasion: (a) it is important to take into account credibility, knowledge, and skill of the source of the feedback is coming from, and (b) this source of self-efficacy is the easiest to undermine and compromise because negative social persuasion can have far greater impact on self-efficacy compared to that of positive persuasion. Unfortunately, individuals who "have been persuaded that they lack capabilities tend to avoid challenging activities that cultivate potentialities and give up quickly in the face of difficulties" (Bandura, 1994, p. 3).

Physiological and affective states. Lastly, self-efficacy can be affected by psychological states. These are such things as stress, fatigue, anxiety, and mood (Bandura, 1994; 1997). If an individual is stressed, or has anxiety, over a particular situation, such as forced to teach biking/cycling with no educational experience, his or her self-efficacy can be negatively impacted. Of the four sources of self-efficacy, physiological states tend to have the least amount of influence on an individual's beliefs. Bandura (1993; 1997), primarily attests this to psychological states often occurring in combination, or even infusion within, master experiences, vicarious experiences, and social persuasion.

Measuring Self-Efficacy Perceptions.

The measurement of self-efficacy perceptions has taken a good deal of criticism in the research community over the past 30 years, much of that criticism stemming from construct validity and measurement problems (Henson, Kogan, & Vacha-Haase, 2001). Bandura (2006: 1997) has refuted nearly all of the initial criticism by publishing a detailed way to avoid such construct and measurement pitfalls.

One of the first issues raised in the criticism of self-efficacy measurement is the use of general self-efficacy instruments that are global in nature and lack a connection to a specific task. As a result, becoming omnibus measures that present problems of predictive relevance and leave ambiguity about what is actually being assessed (Bandura, 1986). During Ross's (1994) examination of teacher self-efficacy perceptions, his meta-analytic study found that nearly all of the 88 studies he examined viewed teacher efficacy as general (global) concept, contrary to the task- and situation-specific designation provided by Bandura (1982; 1997). Ross's findings only help to underline the point that specificity of efficacy measurement has been nowhere near universal in application, and that the construct needs to be continually shifting towards measures that assess precise judgments and are linked to a specific situation or task (Bandura, 1986; Pajaras, 1996).

The literature also supports that teacher self-efficacy is not universal and uniform across all different subject areas. Bandura (1997) made a particular note of this when he stated that teacher efficacy assessment scales, the most common form of assessment in teacher efficacy, should be specific and individualistic to each learning domain. A self-efficacy scale that is tailored towards a mathematics educator should be content specific, with questioning that would be deemed inadequate if asked to a physical educator. "The 'one-measure-fits-all' approach usually has a limited explanatory and predictive value because most of the items in an all-purpose test may have little or no relevance to the domain of functioning" (Bandura, 2006, p. 307). With the emphasis that has been placed on domain specification, researchers looking for predictive or correlational relationship between teacher efficacy and other variables need to implement multifaceted scales that are tailored to the selected activity domain (Bandura, 2006).

> "Efficacy beliefs should be measured in terms of particularized judgments of capability that may carry across realms of activity, under different levels of task demands within a give activity domain, and under different situational

circumstances. Therefore, to achieve explanatory and predictive power, measures of personal efficacy must be tailored to domains of functioning and must represent gradations of task demands within those domains" (Bandura, 1997, p. 42).

Another measurement criticism that has faced self-efficacy is the nature of being a self-reporting measure. Many researchers are skeptical about self-report results that come from questionnaires that ask participants to report on their own perceptions (Feltz, Short, & Sullivan, 2008). However, Bandura (1982) argued that in situations where participants have no reason to distort their responses, self-reporting can be an accurate representation of cognitions. Thus, self-efficacy is best measured when evaluation apprehension has been minimized and participants are asked to respond in private (Feltz et al., 2008).

A further examination of the specific measurement issues and recommended procedures to overcome such problems can be seen in Section IV of this literature review.

Summary of Self-Efficacy

Self-efficacy judgments are content and situation specific, and have been found to be valuable in their ability to influence a person's behaviors, thought patterns, and emotional reactions. These influences primarily result from an individual's mastery experiences, vicarious experiences, social persuasion, and psychological states. Self–efficacy perceptions have emerged in the research community as an effective way to predict behavior and decisions that are made my teachers, and as such, should be further investigated to help understand, and possibly predict, how teachers make instructional choices.

Section III: Methodology for Scale Development

The third section of this literature review will focus on the methodology for self-efficacy scale. This section will be presented in the following segments: (a) introduction to scale development, (b) considering content, (c) generate item pool, (d) determine format of measurement, (e) expert review of items, (f) validity test procedures (e) statistical procedures to analyze items.

Introduction to Scale Development

Over the past several decades hundreds of scales have been developed to measure a countless number of constructs and behaviors (Hinkin, 1995). In research focusing on participant perception, associations, conclusions, and hypothesizes cannot be made until there is an empirical way to quantify these perceptions (DeVellis, 2012). In quantitative research, this is done through the use, or development, of a scale. Perhaps one of the greatest difficulties survey researchers have is assuring the accuracy and validity of the measurement of the construct under examination (Barrett, 1972). To ensure a scale accurately represents the desired construct, researchers are often forced to forgo the use of previously developed scales, and take on the task of developing a new scale that better meets their needs. Unfortunately, it is all too common that newly developed scales are flawed. “The creation of flawed measures may be due in part to the lack of a well-established framework to guide researchers through the various stages of scale development” (Hinkin, 1998, p. 104).

There are a number of imperative steps required in order to design a new instrument or scale. The American Educational Research Association (AERA, 1999) states that an appropriate process for defining a construct with a new scale should include demonstration of content validly, criterion-related validity, and internal consistency.

Considering Content

Under the guidelines of DeVellis (2012), the first step in instrument construction is to determine clearly what it is you want to measure. In survey research, the overall concept that one is trying to measure is known as the construct. Lavrakas (2008), in his *Encyclopedia of Survey Research Methods*, defined a construct as being “the abstract idea, underlying theme, or subject matter that one wishes to measure using survey questions” (p. 133). Constructs can vary significantly, and a great deal of the variability comes from how effectively a primary investigator utilizes relevant literature to define the given construct.

Constructs are defined by the number of variables (also know as dimensions, or facets) they have, with each variable independently acting as a controlling elements towards shaping construct as a whole (Lavrakas, 2008). The level of specificity or generality at which a construct is measured is critical in the initial states of instrument development, and needs to be defined from the onset. (DeVellis, 2012).

Constructs are theoretical abstractions that have definitions represented by two different classifications of meaning: connotation and denotation meanings (Hox, 2012; Bandura 2006). The semantic denotation (operational definition) of the construct is the set of objectives to which the construct refers (i.e., dictionary meaning of a word) (Hox, 2012). A semantic connotation (constitutive definition) of the construct can be described as the "associations it has in the mind of the users, or the list of all characteristics included in the construct" (Hox, p. 56). A connotation, builds upon the denotation, but adds in emotional implications. For example, the denotation for the word *run* would be: to move faster than a walk, never having both feet on the ground at the same time. But the connotation towards that same word can mean different things towards different people, such as: sweaty, fast, hard work, boring, etc.

Both DeVellis (2012) and Hox (2012) summarized that there are three ways that defining a particular theoretical construct can become problematic. The first major issue can stem from connotation issues. Confusion can arise from the construct being associated with more than one meaning, or two constructs point to the same meaning (Hox, 2012, p. 56). Second, complications can come from the denotation, or having a vague understating of what the construct stands for, or actually means. And third, there can also be terminological problems where a label may be applied to a "construct that has wrong characteristics or referents" (Hox, 2012, p. 56; DeVellis, 2012). The literature supports that these problems can be controlled by attempting to: (a) clarifying the connotation (or constitutive) definition of the construct, (b) determining the empirical referents of the construct, and (c) make sure that the label for the construct is understood unequivocally (DeVellis 2012; Hox 2012).

Generate Item Pool

Once the principal investigator (PI) has clearly defined the purpose of the scale, the second step in the scale development process is to generate an item pool. Hinkin (1995; 1998) stated that this process begins by taking a strong theoretical framework and administering a rigorous sorting process that matches a pool of items to the previously defined construct. It is as this point that the researcher seeks to develop items that represent the theoretical construct, thus demonstrating evidence to support construct validity (Bandura, 2006; Hinkin 1998). The domain sampling theory (Ghiselli, Campbell, & Zedeck, 1981) states that it is not possible to measure all aspects of a construct, but that it is important that the sample of items drawn from all potential items sufficiently represents the construct under analysis (Hinkin, 1998).

There are two basic approaches, deductive and inductive, used during the survey item development process (Hinkin, 1995; Hox, 2012). Deductive scale development, also known as 'top-down' analysis, uses a theoretical foundation to provide enough information to create an initial item pool. More specifically, it utilizes previous literature on typology and classification schema of a construct to guide the development of the new survey items. In research scenarios where theory does exit, scale creation using a deductive approach can help assure content validity in relation to the construct (Hinkin, 1998).

Conversely, inductive approach to scale development, also known as 'bottom-up" or exploratory analysis, is used when little theory or literature is available to support the operational definition of a construct (Hinkin, 1995; 1998).

The inductive approach may also be appropriate when "the conceptual basis for a construct may not result in easily identifiable dimensions for which items can then be generated" (Hinkin, 1998, p. 107). Researchers using the inductive approach often start by canvassing respondents, or experts in the field, to provide opinion and recommendations to a line of questioning pertaining to aspects of an abstract construct (Bandura, 2006; Hinkin, 1998; Hox, 2012).

During the initial stages of item pool development, literature has shown that it is effective for researchers to err on the side of over-inclusiveness (Bandura, 2006; Clark & Watson, 1995; DeVellis, 2012; Hinkin, 1998). "The logic underlying this principle is simple: subsequent psychometric analyses can identify weak, unrelated items that should be dropped from the emerging scale, but are powerless to detect content that should have been included, but were not" (Clark & Watson, 1995, p. 311). Both Bandura (2006), and DeVellis (2012), have stated that over inclusion and redundancy is acceptable in the early stages of scale development process. "By using multiple and seemingly redundant items, the content that is common to the item will summate across items while the irrelevant idiosyncrasies will cancel out. Without redundancy, this would be impossible" (DeVellis, 2012, p. 78). Though redundancy may be undesirable in the final instrument, it is much less of an issue in the early stages of item development.

Determine the Format for Measurement

The third step in the scale creation process is to determine the format that is going to be used to measure the construct. Bandura's (2006) guidelines for creating self-efficacy scales emphasize that format should be decided upon early on in the development process. The majority of scales used to measure self-efficacy beliefs, and ones highlighted under Bandura (2006) guidelines, are measured using response scales. "In the standard methodology for measuring self-efficacy beliefs, individuals are presented with items portraying different levels of task demands, and they rate their strength of their belief in their ability to execute the requisite activities" (Bandura, 2006, p. 312). Though Bandura (2006) found there to be variations on the type of response scales used in to measure self-efficacy, the vast majority of them derive from a numerical scale range (0-100 or 0-10), where respondents record their strength of their efficacy beliefs, and assign that belief a numerical rating on the given scale.

Bandura (2006) and DeVellis (2012) both address the fact that scales that use only a few steps should be avoided. Respondents typically avoid extreme positions, thus a scale with only a few step may actually shrink the possibilities to 1 or 2 (Bandura, 2006). Additionally, Bandura (2006) states that having too few point options along a scale creates a lack of; a differentiation that will become much more apparent by the inclusion of intermediate steps along the scale.

Other brief guidelines for scale creation addressed by Bandura (2006), and are supported by Hinkin (1995; 1998), and DeVellis (2012) are:

- "Efficacy scales are unipolar, ranging from 0 to a maximum strength. They do not include negative numbers because a judgment of complete incapability (0) has no lower gradations" (Bandura, 2006, p. 312).
- Content validity is evidenced by phrasing questions in the terms of "can do", rather than "will do". A judgment of a person's ability to perform a task is

based on their perceptions of the skills they have, not their intention to perform the task (Bandura, 2006).

- Statements should be simple and as short as possible, with language that is familiar to the respondent (Bandura, 2006; Hinkin, 1998; Clark & Watson, 1995).
- Items should only address a single issue, with the removal of "double-barreled" questions that have two items being addressed within same question (e.g. I am confident in my abilities to *plan* and *assess* a high school lesson for the volleyball overhand serve) (Hinkin, 1998).

Expert Review of Item Pool

Step four in the instrument development process is to have a group of people who are knowledgeable and have expertise in the content area review the information for construct, content, and face validity. One of the critical reasons that a researcher would use experts for survey instrument review is to confirm or validate the working definition of the construct. Expert reviewers can also be asked to evaluate a scale based on readability, clarity, conciseness, and overall layout. Expert analysis and suggestions are fundamental to the contribution of evidence for content and face validity (Block, Hutzler, Barak, & Klavina, 2013; DeVellis, 2012).

Validity Testing Procedures

One straightforward way of representing how to define validity came from Messick (1995) where he stated "validity is broadly defined as nothing less than an evaluative summary of both the evidence for and the actual (as well as potential) consequences of score interpretation and use (i.e., construct validity conceived comprehensively)" (p. 742). This broad view on validity integrates content, criteria, and consequences as influences on validity to support a construct framework for empirically testing rational hypothesis about score meaning and usefulness (Messick, 1995). Due to the evidentiary support necessary in the validation process it is imperative that validity is also never assumed, and is a continual process of hypothesis generation, data collection an analysis, critical evaluation and logical inference (Downing, 2003).

Though researchers can agree on the overarching theory behind validity, many do not agree on specific evidentiary support necessary to address validity; nor do they agree on the testing procedures to produce such evidence. A major reason for this being the fact that validity is based on a connection between an instrument and a non-observable construct (e.g., self-perceptions), and cannot be exclusively assessed through basic statistical tests, as is the case with reliability (Johnson, 2011). Furthermore, the American Educational Research Association (1999) addressed these disparities in their *Standards for Educational and Psychological Testing* by asserting that "nearly all tests leave out elements that some potential users believe should be measured, and include some elements that some potential users consider inappropriate" (p. 10).

One of the first ways to start addressing evidence necessary during the survey validation process is to specify the sources of validity evidence. Traditionally, there are four sources of validity that are most common in the literature: construct validity, content validity, criterion-related validity, and face validity. Face and content validity are qualitative measures that are often first employed to gain information from expert reviewers based readability, clarity, conciseness, and overall analysis of the survey's representation of the construct (DeVellis, 2012; Block et al., 2013). Following face and content validity procedures, researchers then look to establish the instrument's construct and criterion-related validity prior to testing the instrument on in its first full version. Evidentiary support for both construct and criterion-related validity can be represented by rigorous statistical measures to help ensure the accuracy of instrument construct representation.

Sources of validity influencing survey development.

***Content validity*.** Content validity is a subjective measure of how appropriate elements of an assessment instrument are relevant and representative of the construct of interest (Haynes, Richard, & Kubany, 1995; Messick, 1995). Thus, content validity is directly linked to a well-defined construct prior to being examined, (DeVellis, 2012). For example, there are many variables, or dimensions, that factor in to measuring teacher job satisfaction. If teacher job satisfaction was the construct that one was trying to measure it is imperative that the primary investigator clearly define what they delimit as being factors influencing teacher satisfaction (i.e., items such as salary, benefits, class sizes, relationship with peers, relationship with administration, etc.).

In survey research, evidence to support content validity is built in to an instrument through the item development process (Hinkin, 1995). Survey instrument items that are subject to content validity measures are such things as themes, wording, format, tasks, specific questions, and procedures regarding administration and scoring (AERA, 1999). It is essential that scale item content use the conceptual delimited definition to capture the aspect of the phenomena being measured, primarily excluding additional information that may be outside the researcher's focus for the measure (DeVellis, 2012). Some specific sources of content validity evidence are:

- Representativeness of an instrument to domain
- Instrument specificity
- Matching of item content to instrument specifications
- Logical/empirical relationship of content and domain
- Quality of test questions
- Sensitivity review

(AERA, 1999; Downing, 2003).

The assessment of content validity typically involves field experts who are knowledgeable and have expertise in the content area review the instrument for relevance and focus of the variables of interest. Content review can also be seen as an overall opinion from a on an instrument by a group of highly trained judges (Litwin, 1995). Though content validity evidence is not demonstrated directly through quantified statistics, it provides a "good foundation on which to build methodologically rigorous assessment of survey instrument's validity" (Litwin, 1995, p. 35).

Content validity can often have problems when there is not a clear operational definition of the construct being measured (DeVellis, 2012). This is especially the case when researchers are looking to measure abstract constructs that are not clearly defined in the literature (i.e., the previously given teacher job satisfaction example). Another problem that is often raised when addressing content validity is that it the context of the use is most often non-theoretical, and the practice associated with them is often anecdotal at best (Fulchner, 1999). A third issue that can be seen as problematic with content validity is the ability for researchers to appropriately select representative samples from a desired construct (Fulchner, 1999; DeVellis, 2012). If researchers were seeking to measure our previously discussed construct of teacher job satisfaction, they would define their construct and created sample items representing that definition. If sample items did not incorporate such variables as student teacher ratio or required travel time to work, this abstract construct would be underrepresented and face content validity issues.

Face validity. Face validity is sub-classification of content validity and is established when reviewers examine the instrument to determine if it indeed measures the construct it intends to measure (DeVellis, 2012). Research experts, practicing professionals, or other participants who have prior knowledge pertaining to the construct can be solicited to do this type of review. Of all of the validity procedures, face validity is often seen as the least scientific measure due to the highly subjective nature of results that cannot be quantified. (Lucko & Rojas, 2009). DeVellis (2012) importantly notes that face and content validity are often incorrectly used interchangeably. He clarified this by noting the important differences between the two sources of validity is that content is defined in terms of *specific structured procedures*, and not based on *informal assessments*, such as in purely on face level judgments.

Because subjective judgments, and lack of scientific procedures, act as the basis supporting face validity, the practice has drawn some criticism to its applicability in survey research (Nevo, 1985; DeVellis, 2012; Hinkin, 1998). "The criterion represents a subjective judgment based on a review of the measure itself by one or more experts, and rarely are any empirical approaches used" (Streiner & Norman, 1989, p. 10). Some researchers go as far as not even acknowledging face validity as an appropriate form for measuring validity, or have even placed the phrase in quotation, to de-emphasize its lack of representative weight (Nevo, 1985).

Even with the shortcomings associated with face validity, it is still widely used, and has been seen as an important feature within psychological and educational testing, particularly in the scale development process (Bandura, 2006; DeVellis, 2012; Nevo, 1985). Most notably, Nevo (1985), in his examination of face validity across psychological and education research, concluded that tests with high face validity often have a better change to: (a) include cooperation and positive motivation among participants before and during the test administration, (b) help in attracting and retaining potential participants, and (c) reduce dissatisfaction among participants with low scores.

Criterion-related validity. Criterion-related validity is a measure of agreement between the directly observed results obtained by the survey instrument being tested, with more established results from a similar measurement procedure (Thorndike, 2005). Most often, testing of criterion-related validity correlation is used to describe the relationship between an independent testing measure and the selected criterion; with regression used to for predicting scores in relation to the selected criterion (Hinkin, 1995). The majority of researchers classify criterion-related validity by two different methods: predictive and concurrent validity.

Predictive validity involves using establishing scores from a survey instrument as predictors of the construct being measured, thus forecasting future events, behaviors, actions, and attitude (Bandura, 2006; Litwin, 1995). In order to test for predictive validity, researchers must administer the instrument to a sample with additional measures of an alternative analytical external criteria obtained at a point later in time (Bandura 2006; Haynes, et al., 1995; Messick, 1995). Scores can then be assessed for their ability to predict behavior using correlational and regression analyses.

With concurrent validity, instruments are judged against a well-established instrument (i.e., "the gold standard") or by direct measurement of the same variable (Litwin, 1995; DeVellis, 2012). Essentially, the concept of concurrent validity uses results from the general accepted test as an appropriate benchmark to measure the same construct with a new instrument, thus allowing for correlating between two different sets of measurements (Litwin, 1995; Hinkin, 1998).

The literature supports the fact that criterion-related validity evidence does have the potential to become problematic in the survey development process. Cronbach (1971) addressed one major problem in using criterion-related evidence:

> There is a paradox here. The machinery of validation rests on acceptance of the criterion measure as being perfectly valid

> (save for common error), yet common sense tells one that it is not. Every report of validation against a criterion has to be thought of as carrying the warning clause: "insofar as the criterion is truly representative of the outcome we wish to maximize." (pp. 487-488)

In some research scenarios this can lead to criticism pertaining to the benchmark criterion test being use as possibly even being less valid then the actual independent measurement being tested (Cronbach & Meehl, 1955). Though criterion-related evidence can be useful instrument validation process, the associated criterion can act as a major restrain to detour researchers from providing such evidence (Kane, 1990). Additionally, DeCoster (2006) highlighted the fact that because the creation of a new instrument often derives from a desire to investigate a construct that has not been previously measured, there are often no applicable objective criteria measurements to use as benchmarks. At this point, it is the job of the primary investigator to build a strong argument for a particular interpretation of the new scale by demonstrating that the results are consistent with theoretical variable being measured (DeCoster, 2006; DeVellis, 2012).

Construct Validity. The general consensus in the survey development literature is that the demonstration of construct validation of a measure is the ultimate objective (AERA, 1999; Cronbach & Meehl, 1955; DeVellis, 2012; Downing, 2003). On its most basic level, construct validity is often referred to as an instrument's ability to measure what it is supposed to (Bandura, 2006; DeVellis, 2012). Lavrakas (2008) defined a construct as being "the abstract idea, underlying theme, or subject matter that one wishes to measure using survey questions" (p. 133). Constructs can vary significantly, and a great deal of the variability is associated with how the primary investigator uses the relevant literature to define the construct. Jum Nunnally, in his classic text, *Psychometric Theory* (1978), stated that there are three major aspects to construct validation: (a) identifying the content domain of the construct, (b) empirically examining the extent to which scale items measure that domain, and (c) examining the extent to which the instrument fosters results that have predictability grounded in a theoretical hypothesis. Some specific sources of construct validity evidence are:

- Literature review
- Expert review
- Convergent validity – high relationships with similar measures
- Discriminant validity – low relationships with different measures
- Studies of internal structure – expected relationships among latent traits and items are as expected
- Evaluation of response process – evidence task is eliciting the desired latent trait (think-aloud protocol)
- Response to treatment – test scores change in response to treatment (Groves et al., 2009)

Though there are many different ways to demonstrate evidence representing construct validity, EFA and CFA psychometric testing are most commonly used in the scale development process (Hinkin, 1995). Three different comprehensive reviews (Conway & Huffcutt, 2003; Fabrigar, Wegener, MacCallum, & Strahan, 1999;

Ford, MacCallum, & Tait, 1986) have examined the specific steps necessary in conducting a factor analysis of a new instrument.

Perhaps the most glaring difficulty when measuring construct validity is that constructs are an abstract representation of non-existent observable variable. Additionally, the more abstract the construct, the harder it often is to define and measure (Hinkin, 1998). These abstract constructs often directly influence difficulties with item generation from an exploratory (or inductive) approach, resulting in conceptual consistency being difficult to attain (Hinkin, 1998). Another critical issue that can be problematic in the attainment of construct validity is the inability of researchers to effectively conduct and report EFA and CFA testing and with too much dependency being placed on statistical software (Conway & Huffcutt, 2003; Ford et al., 1986). These authors have also suggested that EFA and CFA procedures should be strongly avoided without adequate training. Furthermore, they attest that this highly technical and difficult training is often hard to find, and a rarity in graduate level programming (Conway & Huffcutt, 2003, Ford et al., 1986).

Threats to validity. All types of research are subject to bias and design threats that can adversely affect validity. There are two different types of threats to validity: internal and external threats (Creswell, 2013). Internal validity relates to how well a study was conducted, and how confident researchers can conclude that changes in the dependent variable was produced exclusively by the independent variables (Leedy & Ormrod, 2001). External validity is related to how well a study's results can be generalized in a different context (i.e., new participants or setting) (Leedy & Ormrod, 2001). External validity threats are most commonly associated with the way respondents are selected, and how they are assigned to groups. Experimental researchers work to identify these threats with the intention to provide enough evidence to limit the threats influence on the overall validity.

There are different threats to validity that can raise numerous questions about a researcher's ability to make conclusions about the construct. Though there are numerous different threats to internal and external validity, examples of threats in typically found in survey research are: (a) inappropriate selection of constructs or measures, (b) insufficient data collected to make valid conclusions, (c) measurement done in too few contexts, (d) measurement done with too few measurement variables, (d) inadequate selection of participants, (e) participants giving biased answers or trying to guess what they should say, (f) invalid experimental method, and (g) not rigorous enough experimental operations (Messick, 1995; Taylor & Asmundson, 2008).

All research studies inherit internal and external validity threats from the outset (Hinkin, 1998). It is the job of the researcher to acknowledge and address validity threats in an open manner, and even the researchers who reject the explicit treatment of these treats are attending to them implicitly (McKinnon, 1988). Just having the knowledge of the potential threats, and the ability to differentiate to what degree they are influencing a study, enable researchers to better analyze results (Taylor & Asmndson, 2008).

In most forms of research, the primary methods used to control for internal and external validity threats come from randomization of participants and sample groups, the use of a theoretically sound research design, statistical analysis that is appropriate to the types of data collected, and the explicit survey items (Groves et al., 2009; Huitt & Hummel, 1999). Additionally, validity threats are controlled when the researcher's peer community generally accepts a study's replicability of methodological processes and results (Groves et al., 2009; Huitt & Hummel, 1999).

Statistical Procedures for Evaluating Items

For survey developers and researchers who wish to measure the empirical evidence supporting construct validity there are a number of statistical analyses that can be performed (Lavrakas, 2008). Factor analysis is the most commonly used analytic technique for data reduction and construct refinement, though many researchers also report inter-item correlation (Hinkin, 1995; Lavrakas, 2008). Factor analyses are procedures used to determine how many latent variables underlie a set of items (DeVellis, 2012). Factor analysis allows researcher to condense information so that variation in original variables can be ideally explained by fewer smaller variables. (DeVellis, 2012).

The two most commonly used forms of psychometric factor analysis are exploratory factor analysis (EFA) and confirmatory factor analysis (CFA). Briefly explained, EFA serves as starting point for measuring factor structure, or interrelationships among items or item groups, of a newly developed instrument, whereas CFA, a more complex approach, is often used subsequently to test a hypothesized factor structure to suggest refinements to improve psychometric properties (Child, 2006; Swisher, Beckstead, & Bebeau, 2004). The consensus among most researchers is that exploratory factor analysis is especially critical in the scale development stages, while confirmatory factor analysis is preferred when measurements have a well-defined underlying theory for hypothesized patterns for loading factors (DeVellis, 2012). The factor analysis of the PEES-LPA instrument will use exploratory factor analysis due to the construct of teacher self-efficacy towards instruction of lifetime physical activity not being clearly defined, thus ruling out the any reasoning to use confirmatory factor analysis.

Sample size. Exploratory factor analysis (EFA) is a statistical technique used to examine the possible underlying factor structure of a set of observed variables, and is often used by researchers during the early stages of scale construction (Burton & Mazerolle, 2011). In other words, EFA takes a large group of items and looks for a way that the data may be reduced, or summarized, using fewer items (Lamarche, Gammage, Sullivan, & Gabiel, 2013). This factor analysis process has been found to be highly dependent on sample size, and much of the literature supports that as sample size increases the likelihood of statistical significance increases (DeVellis, 2012; Thorndike, 1997; Hinkin, 1995).

Currently, there is no general consensus (or rule) for the required sample size to conduct factor analysis. Much of the literature supports that factor patterns emerging from large-sample factor analysis are usually more stable than emerging from small samples (DeVellis, 2012; Hinkin 1995). The question of "how large is large enough?" has generated a wide array of responses from survey researchers. Suggestion range from the following:

- Five to 10 participants per item, up to about 300 (Tinsley & Tinsley, 1987)

- Up to 100 participants is poor, up to 200 is fair, up to 300 is good, up to 500 is very good, and up to 1000 is excellent (Tinsley & Tinsley, 1987)
- 150 participants for EFA (Guadagnoli & Velicer, 1988)
- Costello and Osborne (2005) examined "best practices" in EFA and found that the majority of the studies they reviewed to have a subject-to-item ration of 10:1
- 200 participants are adequate for factor analyses that involve no more than 40 items (Comrey, 1988).

Though much of the literature leans towards suggesting larger sample sizes constitute better EFA, there is still no definitive empirical evidence. Some researchers have refuted the sentiment of large sample sizes by countered that when factor loadings are high (i.e., 0.6 – 0.8), fewer subjects can be adequate ($n \geq 100$) (Field, 2009). The PEES-LPA will require 200 responding participants as its minimal sample size.

Factor extraction model. Principal axis factoring (PAF) is debatably the most widely used method of factor extraction when EFA is the chosen analytic method (Conway & Huffcutt, 2003). The primary reason is this estimation method requires researchers to estimate the amount of communality by running one variable in question (e.g., item #1) as the outcome and all other items (items #2 and above) as predictors (Field, 2009). Once these are completed, new commonalities can be calculated that represent the correlation between the variable and the extracted factors (Field, 2009).

The subsequent steps to examine the construct are:

1. Test multivariate assumptions
2. Consider rotation method. Factor rotation alters the patterns of the resultant factor loadings from the PAF, which allows for improvement in interpretation. There are two types of rotation: orthogonal and oblique. Orthogonal rotation there is no correlation between extracted factors, with oblique there is. Field (2009) states that "the choice of the rotation depends on whether there is a good theoretical reason to suppose the factors should be related or independent" (p. 439)
3. Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy & Bartlett's test for sphericity. The KMO and Bartlett's test both test the suitability of the respondent data for factor analysis.
4. Cumulative percentage of variance. Test that examines the overall variance that is explained by all items; should have an eigenvalue greater than 1.0 to be retained. (DeVellis, 2012).
5. Scree test. Graphical representation of the independent weighted combinations of the original variables with corresponding factors. (Field, 2009).
6. Parallel analysis. Used to measure how many factors to extract or retain (Field, 2009; DeVellis, 2012). Uses actual eigenvalues are compared to random order eigenvalues. Factors are retained when actual eigenvalues surpass random order ones (Field, 2009).
7. Test for internal consistency reliability

Summary of Scale Development

One overall inference drawn from the literature when testing a measure for evidence to support validity; more appears to be better. The more different sources of validity can be tested and re-tested, the additional information that researchers will have addressing the quality of the survey's ability to measure the given construct. To the best of the researcher's ability, using both statistical and non-statistical testing from all four sources of validity is advantageous. Specifically relating to survey research, Bandura (2006) focused on the minimal standards in survey validity testing. Using Bandura's (2006) benchmarks, as well as the information from the provided literature, a summation of procedures most relevant in validity testing are: (a) pretesting items and many times as possible, (b) factor analysis to verify the homogeneity of items, (c) expert review testing the instrument's ability to measure the defined construct, and (d) testing for both predictive and discriminative validity. Though each of these validity procedures has specific, and often complex, sub-procedures, the general findings are relevant in serving as the initial foundation for validity procedures for construction and testing of a new instrument: the PEES-LPA.

Section IV: Measurement of Physical Education Teacher Efficacy

The literature has shown that there is limited research pertaining to self-efficacy perceptions towards instruction in physical education. Whether this is due to the time necessary to create and validate an instrument, the lack of value placed on physical education, or physical education not being a core subject, the gap in the researcher is present. The focus of this section was to examine the current self-efficacy instruments specific to physical education teacher perceptions.

Though physical educator self-efficacy perceptions have been examined using general teacher efficacy scales, this section will be restricted to focus on 11 instruments that were built specifically aimed at self-efficacy perceptions of physical educators. The 11 instruments sampled were evaluated based on: (a) how researchers define and measure constructs, (b) procedures used during sampling and testing, as well as (c) reliability and validity methods employed. This section will be broken down into the following sections: (a) introduction, (b) date of publication, (c) content, (d) format, (e) validity procedures, and (f) reliability procedures.

Introduction

There is a great deal of literature to support the fact that there is pronounced variation in the breadth and depth to validity and reliability testing in survey design. A major take-away from the literature is that there is no definitive benchmark that a survey instrument must meet when providing evidence to support validity and reliability. Though not universally agreed upon, numerous researches have stated that there are methodological steps that can be taken to influence the applicability of psychometric procedures being performed (Bandura 1997, 2006; DeVellis, 2012; Henson, et al., 2001). To support this point DeVellis (2012) noted that "psychometrics clearly have emerged as a methodological paradigm in its own right" (p. 8). He later goes on to address the impact of psychometric evaluation for scale development:

1. There is a widespread use of psychometric definitions of reliability and validity
2. The popularity of factor analysis in social science
3. The adoption of psychometric methods for developing scales measuring and array of variables far broader than those with which psychometrics were initially concerned

(DeVellis, 2012, p. 8)

Although numerous measurement scales have been created to explore the self-efficacy beliefs of classroom teachers, currently there are very few scales that are specific to physical education. To support this point, Martin and Kulinna (2005) found that there was very little research examining the self-efficacy beliefs of physical educators, and that much of the research that was present was far too broad, with little examination placed on specific pedagogical decisions made by teachers. This statement still holds true in 2013.

Inclusion of instruments for this section was based on the fulfillment of the following criteria: (a) the instruments were specific the self-perception measurement of self-efficacy of physical educators, (b) the instruments only targeted self-efficacy perceptions of teachers, not students, (c) the study reported methods and results for evaluating reliability and validity from its original source (no replication of previously published instrument) (see Table 2).

Date of Publication

First, all of the self-efficacy instruments had quite recent publication dates. The publication date ranges for the instruments found were from 1997 through

2013, with a mean year of publication being 2008. The literature supports that prior to 1997, physical educators were assessed on their self-efficacy perceptions, but none of the survey instruments used to measure those perceptions were context specific towards teaching physical education. Many instruments used prior to 1997, and even still used today, are general teacher efficacy scales (i.e., the *Teacher Efficacy Scale,* by Gibson and Dembo [1984]), with no specificity to any subject or content area. This is an area of concern due to the Bandura's (1977; 1982; 1986) specification for efficacy measurements being subject and content specific.

This point is especially strengthened with the fact that of the 8 out of the 11 (73%) survey instruments that were published in peer-reviewed journals, with a mean publication year of 2008 (SD = 5.2). Though the reasoning for lack of publication is unknown for the 3 instruments linked to dissertations & masters' thesis, the inferences drawn from these publications do not hold the same weight compared to peer-reviewed studies. For the purposes of this examination, the non-peer reviewed instruments will be included in the discussion due to their reporting of a formalized procedure for testing for validity and reliability evidence (see Table 2).

Content

Scales were evaluated based on their measurement of physical education teacher self-efficacy perceptions using an instrument not previously published. Of the 11 instruments found, six groupings appeared:

1. General PE teacher efficacy (n = 2)
2. Efficacy towards APE instruction & inclusion (n = 4)

3. Efficacy towards teaching high school physical education (*n* = 1)
4. Efficacy towards teaching physically active classes (*n* = 2)
5. Efficacy towards teaching obese students in physical education (*n* = 1)
6. Other efficacy measures specific to physical education instruction (*n* =1)

One major concern arises from this data. Primarily, if self-efficacy is such a well-supported construct in educational research, why do the 11 instrument sampled represent such a large portion of the self-efficacy research in physical education? The answer comes back to the thoughts of Martin and Kulinna (2005) addressed that the scale development process is time consuming, and that self-efficacy research has not been accurately explored in physical education (see Table 2).

Format of Measurement

It is also clear from the descriptive results that unipolar Likert scales are the primary vehicle for measuring physical education teacher perceptions of self-efficacy; though not all researchers are in agreement as to the number of scale points that are appropriate. All efficacy instruments coded used a unipolar Likert scale for measuring self-efficacy perception, though there was a great deal of variance on how this Likert scale was administered. Likert scales ranged from a 6-point scale (1-5) to an 11-point scale (0-10). Bandura (2006) noted that instruments that use very few scale points, such as a 3 (1-3) or 4 (1-4), should be avoided due to the fact that many survey takers typically like to avoid the extreme response positions. "Including too few steps loses differentiating information because people who use the same response category may differ if intermediate steps were included" (Bandura, 2006, p. 312). There was also discrepancy in the total number of survey items on the instrument. The range of total survey items on each instrument was between 10 and 53, with the mean number of survey items being 24.8 (SD = 13.16). According to the recommendations from Bandura (2006) and DeVellis (2012), all of the 11 instruments in this study met the criteria for scale response points.

Though the creation of the PEES-LPA will be conducted under the recommendations of Bandura (2006), it is important to note that not all measurement theorists agree with his recommendations for Likert scale development (Henson, et al., 2001). Clark and Watson (1995) stated that "it must be emphasized also that providing more response alternatives (e.g., 9-point scale versus a 5-point scale) does not necessarily enhance reliability or validity" (p. 313). In fact, increasing the number of alternatives may actually decrease the validity if respondents are unable to make the more subtle distinctions that are required. Additionally, having too many alternatives can produce an issue where scores are rendered less valid (Clark & Watson, 1995)(see Table 2).

Sample Size

The final descriptive item examined was associated with the sample size reported of the instrument's final version on in-service or pre-service teachers (post expert review and pilot phase). The sample sizes ranged from 50 to 357 participants with a mean sample size of 264.7 (SD =176.82). This finding is especially alarming because 5 of the 11 instruments (45%) fell below the minimum 200 sample size that is recommended in much of the survey development literature.

Table 2
List of Physical Educator Self-Efficacy Instruments Selected for Review

Scale Name	Authors	Instrument Focus	Year of Publication
1. Teacher's Self-Efficacy Scale for High School Physical Education Teachers (TSES-HSPET)	Pan, Y.	Efficacy towards teaching high school physical education	2012
2. Physical Education Teaching Efficacy Scale (PETES)	Humphries, C.A, Hebert, E., Daigle, K., Martin, J.	General PE teacher efficacy	2012
3. Self-Efficacy Instrument for Physical Education Teacher Education Majors Towards Inclusion (SE-PETE-D)	Block, M.E., Hutzler, Y.S., Barak, S., Klavina, A.	GPE teacher efficacy towards inclusion in general PE classes	2013
4. Physical Education Teachers Physical Activity Self-Efficacy Scale (PETPAS)	Martin, J.J. Kulinna, P.H.	Efficacy towards teaching physically active classes	2003
5. Adapted Physical Education Teacher Self Efficacy Scale (APE-TES)	Haeusner, M.K.	General APE teacher efficacy	1997 (Dissertation)
6. General Physical Education Efficacy Survey (GPEES)	Umhoeffer, D.L.	GPE teacher efficacy towards delivery of APE services	2011 (Thesis)
7. Physical Educator's Self-Efficacy Toward Including Students with Disabilities (PESEISD-A)	Taliaferro, A.	GPE teacher efficacy towards inclusion in general PE classes	2010 (Dissertation)
8. Self-Efficacy Theory Measures (SETM)	Gorozidis, G., Papaioanno, A.	Efficacy towards: (1) teaching lesson plans, (2) various teaching styles, and (3) promoting students' exercise	2011

Validity Procedures

The validity procedures examined were exploratory factor analysis, confirmatory factor analysis, expert review, and factor load ranges. All of these validity procedures helped to measure if in fact the survey measures what it is indented to measure (American Educational Research Association, 1999; DeVellis, 2012). The validity measures used across the different scales showed a great deal of discrepancy. From the coded instruments selected, 73% of the studies used exploratory factor analysis, and 55% of the studies used confirmatory factor analysis. These findings are directly in line with much of the literature to associated factor analysis during scale development. Hurley et al. (1997) underlined this point by affirming that though exploratory and confirmatory factor analysis both seek to explore the variance associated with a set of variables with a smaller set of latent variables, exploratory factor analysis is often more appropriate during the scale development process. The consensus among most researchers is that exploratory factor analysis is especially critical in the scale development stages, while confirmatory factor analysis is preferred when measurements have a well-defined underlying theory for hypothesized patterns for loading factors (DeVellis, 2012).

Unfortunately, findings showed that only 55% of the selected efficacy instruments demonstrated the use of expert review. This is especially concerning due to the fact that this step in scale development process has been found to be critical in (a) having experts confirm or invalidate your working phenomena, (b) evaluate item relevancy, (c) evaluate item clarity and conciseness, and (d) point out ways that you have missed out on measuring the phenomena (Hinkin, 1995;

DeVellis, 2012). The instruments from this analysis that did not make use of expert review leave them subject to questioning over the accuracy to meet content and construct validity.

One example of validity inconsistencies seen in the instruments was the use of double-barreled questioning. An example of this can be seen in the *Physical Education Teacher Efficacy Scale* (PETES) created by Humphries, et al. (2012), raised concerns pertaining to construct validity and item creation, primarily due to the use of some double barreled questions. An extracted example from the PETES questionnaire demonstrating double-barreled questioning are:

PETES Item #13: I know what the NASPE standards are, and can plan and teach them accordingly.

PETES example item 13 clearly demonstrate the use of double-barreled questioning. In item number 13, most physical educators would see planning and teaching to be two very different teaching task demands, and some may even argue that the construct 'to teach' is also not clarified enough to accurately measure teacher perceptions. These double-barreled questions go strictly against the guidelines of self-perception research (Bandura, 2006; Humphries et al., 2012), and should be re-examined before replication in future research.

Instruments found with strong psychometric results demonstrated that they go far beyond the specified criteria for psychometric testing delimited in this review. Additional exploratory psychometric testing utilized demonstrated by these instruments were: Kaiser-Meyer-Olkin (KMO) index of sampling adequacy, Bartlett's test for sphericity, variance inflation factor (VIF) of mulitcollinearity, and crossoloading. Additional confirmatory psychometric testing utilized were: normal fit index (NFI), root-mean-square error of approximation (RMSEA), and comparative fit index (CFI) (Block et al., 2013; Humphries et al., 2012) (See Table 3).

Reliability Procedures

The reliability measure that was most widely used to evaluate efficacy instruments was Cronbach's alpha test for internal consistency (91%). The alpha level of .70 has been widely noted as being acceptable in survey research (Cronbach, 1951; Bandura, 2006), and thus was demarcated as the minimal acceptance level for this review. Most of the articles (91%) reported alpha levels for all dimensional factors being greater than .70, with one article (Biddle & Goudas, 1998) reporting a single alpha level to represent the entire scale. A trend that was detected in this review, and supported by the literature for scale development and analysis, was that items typically reducing the alpha level of a particular factor were often eliminated in the pilot or test phases to keep with the .70 acceptable value. It was also concerning that only 4 out of 11 studies (36%) used test-retest as measure to support further reliability evidence. More specifically, of the four studies reporting test-retest measurement evaluation, only two of those four were from peer-reviewed journals (see Table 3).

Conclusions of Measurement in Physical Education

One overall inference drawn from this analysis pertaining to reliability and validity testing during instrument construction is that generally speaking, more appears to be better. The more constructs, content, and external validity can be tested and re-tested, the more information that researchers will have linking the quality of the survey's ability to measure the given construct. Testing for internal consistency (Cronbach's alpha) appears to be the minimal benchmark a quality survey analysis should undergo prior to completion. Finally, under the same 'more is better' principle, a greater sample size with the instrument in its final version has been linked with the strongest results for both reliability and validity.

Part V: Conclusion

The major goal of the scale development process is to create an instrument that is valid in measuring a specified construct. Validity is a term that has somewhat different meanings, even within survey research, and can vary from researcher to researcher (Groves et al., 2013). Specifically in survey construction, validity refers to an instruments' ability to measure the construct it is intended to measure by presenting sources of evidence to support or refute the meaning or interpretation of the test scores from that measure (DeVellis, 2012).

Determining whether a scale provides enough evidence to support validity has been an area of disagreement among researchers for the past few decades (AERA, 1999; Hinkin, 1995). Even with the disparities among researchers specific to scale validation procedures, all are in agreement to the fact that validity is an absolute necessity during the process of scale development, and without evidence of validity, measurements hold little-to-no intrinsic meaning (Downing, 2003).

Due to the clear focus placed on teaching lifetime physical activities in secondary physical education, it has become apparent that there is a need for researchers analyze the how these types of lifetime activities are being taught in schools, and what teachers specifically think about their abilities to teach these activities. Although numerous measurement scales have been created to explore the self-efficacy beliefs of classroom teachers, currently there are very few scales that are specific to physical education. To support this point, Martin and Kulinna (2006) found that there was very little research examining the self-efficacy beliefs of physical educators, and that much of the research that was present was far too broad, with little examination placed on specific pedagogical decisions made by teachers (Martin & Kulinna, 2006).

Table 3

Validity and Reliability Procedures Reported for Physical Educator Self-Efficacy Instruments

				Meets Delimited Requirements?										
				Identification and Characteristics			Validity				Reliability			Total Number of Delimited Requirements Met (out of 10)
			Year of Pub.	Peer Reviewed	Sample Size (n>200)	# of Scale Points (5-11 per item)	Expert Review Conducted	EFA Conducted	CFA Conducted	Factor Loading (loadings>.40)	Cronbach's Alpha (α >.70)	Chi Squared Conducted	Test-Retest Conducted	
1	Teacher's self-efficacy scale for high school physical education teachers (TSES-HSPET)	Pan, Y.	2012	✓	✓	✓		✓	✓	✓	✓	✓		8
2	Physical education teaching efficacy scale (PETES)	Humphries, C.A., Hebert, E., Daigle, K., Martin, J.	2012	✓	✓	✓	✓	✓	✓	✓	✓	✓		9
3	Self-Efficacy instrument for physical education teacher education majors towards inclusion (SE-PETE-D)	Block, M.E., Hutzler, Y.S., Barak, S., Klavina, A.	2013	✓	✓	✓	✓	✓	✓	✓	✓	✓		9
4	Physical Education Teachers Physical Activity Self-Efficacy Scale (PETPAS)	Martin, J.J., Kulinna, P.H.	2003	✓	✓	✓		✓	✓	✓	✓			7
5	Adapted Physical Education Teacher Efficacy Scale (APE-TES)	Haeusner, M.K.	1997			✓	✓	✓		✓	✓			5
6	General Physical Education Efficacy Survey (GPEES)	Umhoefer, D.L.	2011			✓	✓			(Not Reported)	(Not Reported)		✓	3
7	Physical Educators's Self-Efficacy Toward Including Students with Disabilities (PESEISD-A)	Taliaferro, A.	2010		✓	✓	✓	✓	✓	✓	✓		✓	8
8	Self Efficacy Theory Measures (SETM)	Gorozdis, G., Papaioannou, A	2011	✓	✓	✓		✓	✓	✓	✓			7
9	Self-Efficacy Questionnaire of Physical Education Teachers with Respect to Obese Students (SEQPET-OS)	Martinez-Lopez, E., Sanchez, M., Alvarez, M., Cruz, M.	2010	✓		✓		✓		✓	✓			5
10	Exemplar Physical Education self-efficacy scale (EPEC-ES)	Martin, J.J., McCaughtry, N., Hodges-Kulinna, P., Cothran, D.	2008	✓		✓	✓			(Not Reported)	✓		✓	5
11	Physical Education Teacher Efficacy Scale (PETE)	Biddle & Goudas	1998	✓		✓				(Not Reported)	✓		✓	4
		Column Total		8	6	11	6	8	6	8	10	3	4	

CHAPTER 3

METHODOLOGY

The purpose of this study was to develop and begin the validation process of an instrument that measures efficacy perceptions of physical educators towards teaching lifetime physical activities. This instrument, the *Physical Educator Efficacy Scale for Teaching Lifetime Physical Activities* (PEES-LPA), was developed through expert review and numerous pilot procedures based on Bandura's Self Efficacy Theory (1977; 1982). The PEES-LPA was constructed and validated using the recommendations and guidelines presented in the previously mentioned literature, paired with Bandura's *Guide for Constructing Self-Efficacy Scales* (2006), and DeVillis's *Scale Development: Theory and Practice* (2012).

This study was conducted in four different phases: (a) Phase I- Item generation, (b) Phase II- Prepilot review, (c) Phase III, Validation study, (d) Phase IV, Assessment of reliability and construct validity (Scrabis-Fletcher & Silberman, 2010). This chapter will outline the steps in each of these four phases.

Phase I: Item Generation

The steps in Phase I were developed based on the supported literature on item generation and initial content and construct validity (Bandura, 2006; DeVellis, 2012; Hinkin 1995, 1998; Krosnick & Fabrigar, 1997). Phase I consisted of the following five steps: (a) approval from institutional review board, (b) developing construct content, (c) lifetime physical activity content, (d) item pool generation, and (e) format.

Approval from Institutional Review Board

This phase of the study involved obtaining approval from the Institutional Review Board (IRB) at the University of Virginia, in Charlottesville, Virginia. The following protocol, number 2014-0020-00, was reviewed and approved by the IRB on February 12, 2014 (see Appendix A).

Developing Construct Content

When considering content for a new self-efficacy instrument scale both Bandura (2006) and DeVellis (2012) stressed the importance of having a construct that can clearly be defined, and that is distinguishably different from measurements addressing other aspects of self-perception (e.g. self-esteem, locus of control, and perceived competence). The construct measured in this study was physical education teacher efficacy towards teaching lifetime physical activities.

Constructs developed under Bandura's self-efficacy umbrella must also specifically address both task and situation specific perceptions of confidence (Bandura, 1997). To align the PEES-LPA properly with self-efficacy theory, task specificity was addressed through teaching behaviors and guidelines presented by NASPE's (2008) *National Initial Physical Education Teacher Education Standards*, and situation specificity was addressed though the use of selected lifetime physical activity categories.

Lifetime Physical Activity Content

In the fall of 2013, the American Alliance for Health, Physical Education, Recreation, and Dance (AAHPERD) published an extensive chart detailing grade-specific outcome expectancies in K-12 physical education (AAHPERD, 2013). Within this publication, AAHPERD provided an operational definition of activity categories taught in physical education, listed specific activities within those categories, as well as designated which of the activities meet the requirements for being a lifetime physical activity. The activity categories assigned by AAHPERD (2013) are:

- Outdoor Pursuits (e.g., hiking, backpacking, snow/water skiing, mountain biking).
- Fitness Activities (e.g., running, cycling/biking, yoga, weight/resistance training).
- Dance & Rhythmic Activities (e.g., ballet, modern, line, social and square)
- Aquatics (e.g., swimming, diving, water aerobics)
- Individual Performance Activities (gymnastics, track and field, self-defense)
- Games & Sports (e.g., invasion games/sports, target & fielding games/sports, net & wall games/sports)

  (see Table 4)

Table 4

AAHPERD Operational Definition of Activity Categories

Outdoor Pursuits		
Hiking	Backpacking	Fishing
Orienteering/Geocaching	Ice Skating	Skateboarding
Snow or Water Skiing	Snowboarding	Snowshoeing
Surfing	Bouldering/Climbing	Mountain Biking
Adventure Activities	Ropes Courses	
Recreational boating (e.g., kayaking, canoeing, sailing, rowing)		

Fitness Activities		
Yoga	Pilates	Cycling/Biking
Spinning	Running	Fitness Walking
Kickboxing	Cardio-kick	Zumba
Exergaming	Weight/Resistance Training	

Dance & Rhythmic Activities		
Creative movement/Dance	Ballet	Modern
Ethnic/Folk	Hip Hop	Latin
Line	Ballroom	Social and Square

Aquatics		
Swimming	Diving	Water Polo
Synchronized Swimming	Water Aerobics	

Individual Performance Activities		
Gymnastics	Figure Skating	Track & Field
Multi-Sport Events	In-line Skating	Wrestling
Self-Defense		

Games & Sports	
Invasion Games/Sports	Target & Fielding/Striking Sports
Net/Wall Sports	

Lifetime Activities

- Includes any of the above-mentioned categories: outdoor pursuits, selected individual performance activities, aquatics, net/wall sports, and target sports.

- Note: Inclusion and fielding/striking games were excluded from the secondary physical education outcomes due to the fact that the activities require team participation and are less suited to lifelong participation (AAHPERD, 2013, p. 39).

  (AAHPERD, 2013)

AAHPERD (2013) defined that *lifetime physical activities* represent the categories of: (a) outdoor pursuits, (b) selected individual performance activities, (c) aquatics, (d) net/wall sports, (e) target games, (f) dance and rhythmic activities, and (g) fitness activities with specific exclusion of invasion games/sports, fielding & striking games/sports. Due to the minimal literature on specific activities that constitute being a *lifetime physical activity*, the Principal Investigator (PI) used the AAHPERD categories (with allocated activities) as the content framework driving the elicitation study.

AAHPERD, the national governing body for physical education in the United States, created the physical activity areas through a curriculum framework task force comprised of field experts and leading researchers around the country. Using the AAHPERD (2013) recommendations as the platform for this portion of the study, the classification of a *lifetime physical activity* was delimited for the Elicitation Study by the following criteria:

1. *Lifetime physical activities* represent the categories of outdoor pursuits, selected individual performance activities, aquatics, net/wall sports, & target games (AAHPERD, 2013).
2. Invasion games and fielding/striking games were excluded because they require team participation and are not appropriate for lifelong participation (AAHPERD, 2013).
3. To address the most wide-spread lifetime physical activities taught in the United States, activities were eliminated based on their geographical specification (e.g., ice skating, surfing, bouldering/climbing)

4. Net/Wall games were expanded to represent: badminton, table tennis, tennis, racquetball/squash, and pickleball
5. Target games were expanded to represent: archery, croquet, golf, horseshoes, bocce, bowling, and disc golf.

A final list of delimited lifetime physical activities that was prepared for use in the elicitation study can be seen in Table 5.

Item Pool Generation

The next step was to generate a set of items that fully represent the conceptual domain of the construct: self-efficacy perceptions of physical educators towards instruction of lifetime physical activities (Mackenzie, Podsakoff, & Podsakoff, 2011). This large pool of items acted as candidates for eventual construction of the final instrument for both the pilot and validation studies (DeVellis, 2012).

A deductive approach was used to develop task constraints that influence self-efficacy perceptions towards instruction of lifetime physical activities. The PI selected two publications, which were both nationally recognized, and expert reviewed, to guide the item generation of teaching impediments or obstacles that influence self-efficacy: (a) *National Initial Standards for Physical Education Teacher Education* (NASPE, 2008; NCATE, 2008), and the (b) *Physical Education Teacher Evaluation Tool* (NASPE, 2007). The *Standards* and the *Evaluation Tool* (NASPE, 2008; 2007) addressed teaching task constraints in six different areas: scientific and theoretical knowledge, skill and fitness based competence, planning and

Table 5

List of Lifetime Physical Activities Selected for PEES-LPA

Outdoor Pursuits		
Hiking	Backpacking	Fishing
Orienteering/Geocaching	Mountain Biking	Ropes Courses/Climbing
Fitness Activities		
Yoga	Pilates	Cycling/Biking
Spinning	Running	Fitness Walking
Kickboxing	Zumba	Exergaming
Weight/Strength Training		
Dance & Rhythmic Activities		
Creative movement/Dance	Ballet	Modern
Ethnic/Folk	Hip Hop	Latin
Line	Ballroom	Social and Square
Aquatics		
Swimming	Diving	Water Aerobics
Individual Performance Activities		
Gymnastics	Track & Field	Multi-Sport Events
In-line Skating	Self-Defense	
Net/Wall Games		
Badminton	Tennis	Table Tennis
Racquetball/Squash	Pickleball	
Target Games		
Archery	Golf	Bowling
Disc Golf	Croquet	Horseshoes
Bocce		

Note. Adapted from American Alliance for Health, Physical Education, Recreation, & Dance-AAHPERD (2013). *National standards & grade level outcomes in K-12 physical education*. Reston, VA; AAHPERD.

implementing, instructional delivery, impact on student learning, and professionalism.

Example NASPE standards (2008):

- 2.3. Demonstrate performance concepts related to skillful movement in a variety of physical activities.
- 3.5. Plan and adapt instruction for diverse student needs, adding specific accommodations and/or modifications for student exceptionalities.
- 4.3. Provide effective instructional feedback for skill acquisition, student learning and motivation.

During this phase of item generation, construct-formation involved and narrowing the *Standards* and *Evaluation Tool* (NASPE 2008; 2007) in to empirical referents, or criteria used to measure and evaluate the absence or presence of the defining attributes, for the presently undefined term 'efficacy to teach' (Hox, 2012). The PI, under advisement from members of the Kinesiology faculty at the University of Virginia, reviewed each of the two NASPE publications to generate a list of items that represented teaching task constraints associated with instruction of lifetime physical activities. Prior to the prepilot review, the final list consisted of perception questions (*N* = 8) addressing the following areas:

1. Perception of personal skill ability for lifetime physical activity category
2. Participating in each lifetime physical activity category over the course of the past year
3. Competence to teach a wide variety of skills in each activity category compared to the average physical education teacher

4. For each lifetime physical activity category, Confidence in ability to:
    a. Identify and define critical skill elements
    b. Present critical skill elements (describe & demonstrate)
    c. Provide skill-specific feedback for critical skill elements
    d. Accurately assessing student learning
    e. Adapt instruction for diverse student needs

Format

The Pre-Pilot questionnaire consisted of three different sections: (a) demographic information, (b) items measuring perception towards personal ability, and (c) items measuring efficacy towards instruction of lifetime physical activities.

The first section addressed demographic information (*n* = 5) for the participants. These demographic items were represented by both drop-down multiple-choice questions and open-ended questions. The drop-down multiple-choice questions (*n* = 4) were: grade level currently teaching (HS/MS/Both), gender (male/female), state currently teaching (50 US states), highest educational degree obtained (bachelors/masters/doctorate). The open-ended question (*n* = 1) was: number of fulltime years of physical education teaching experience (numerical value).

The second section of the questionnaire consisted of three perception items addressing the participant's experience with lifetime physical activities. Inclusion of these items represented literature supporting self-efficacy perceptions being highly influenced by mastery experiences. (Bandura, 1997; 1982; Tschannen-Moran et al, 1998; Tschannen-Moran & Woolfolk-Hoy, 2001). Additionally, there was literature to support that controlling for talent and opportunity can highlight the significance self-efficacy perceptions have on successful achievement in many aspects of human functioning (Bandura, 1997; Caprara, et al., 2003; Maddux, 1995). These perception items were presented in an 6-unit (0-5) scale with the following qualitative label descriptors: no experience, novice, advanced beginner, competent, proficient, expert (Malinen, Savolainen, & Xu, 2013).

The third section of the questionnaire consisted of six items addressing efficacy perceptions of explicit teaching behaviors during instruction in physical education. These items included wording such as: (a) "How confident are you in your abilities to provide skill specific feedback of critical skill elements for *Outdoor Pursuits* activities….", and (b) "how confident are you in your ability to adapt instruction for diverse student needs for *Outdoor Pursuits* activities…". Based on the recommendations of Bandura (2006), an 11-unit (0-10) response scale was used with the following qualitative descriptors: 0 = *no confidence*, 5 = *moderate confidence*, and 10 = *complete confidence*.

In total, the Pre-Pilot questionnaire consisted of five demographic items and 56 self-efficacy items. The PI, and members of the Kinesiology faculty at the University of Virginia, reviewed the items for content validity, item inclusion, clarity, and conciseness.

Phase II: Pre-Pilot Review

The steps in Phase II were developed based on the supported literature (Bandura, 2006; DeVellis, 2012) and consisted of two steps: (a) focus group review and (b) expert review.

Table 6

Summary of Steps in Phase I

Step	Purpose	Source
1. Approval from IRB Board	• Federal regulations require that all proposed human research studies undergo review by the Institutional Review Board (IRB). • The primary role of the IRB is to protect the safety and welfare of human subjects.	UVA IRB for Social and Behavioral Sciences
2. Developing Construct Content	• Clearly define and set boundaries on selected construct	Bandura (2006) DeVellis (2012)
3. LPA Content Selection	• Defining LPA categories • Defining activities that fall within each category	AAHPERD (2013)
4. Item Pool Generation	• Generate a set of items that represent construct • Items pair 7 LPA categories with 6 teaching task constraints that most influence efficacy perceptions towards instruction	AAHPERD (2013) DeVellis (2012) NASPE (2008) Hox (1997) Hinkin (1995) Bandura (2006)
5. Formatting	• Create three different sections: demographic items, perception towards ability items, efficacy towards teaching LPA items	Bandura (2006)

Focus Group Review

The Pilot instrument was then blindly administered to a convenience sample of doctoral students (*N* = 5) in Kinesiology. The focus group was asked to: (a) take the complete survey and then review and make comments on the instrument's (b) content, (c) readability, (d) clarity, (e) format, and (f) length. The focus group provided further evidence of content validity by evaluating the extent to which the items in the survey items accurately represented the construct, and if teaching behaviors and tasks presented were appropriate in the domain of interest. Every item of qualitative feedback was recorded, and taken under considered by the PI and his advisor for future revision. Revisions were made on the items and format after the discussion with the focus group to create the first full version of the instrument prior to pilot testing. Accepted changes were used to revise the scale for item retention, item face and content validity, confirm or invalidate definition of the phenomena, and point out ways of tapping in to the phenomena that may have failed to be included (DeVellis, 2012).

Expert Review

During the next phase, physical education teacher education experts reviewed the Pilot instrument (Bandura, 2006; DeVellis, 2012). Arguably, the involvement of domain experts to review subject matter before, during, and after is the strongest evidence for construct, content, and face validity (Lucko & Rojas, 2010). The first draft of the instrument was distributed to a group of experts in the field of physical education pedagogy (*N* = 11).

Recruitment. Experts were physical education pedagogy professors at doctoral-granting universities across the United States. These experts were selected based on the following criteria: (a) they were identified in the literature as experts in research in physical education, (b) were former pedagogy or research methods professors of the PI, or (c) who demonstrated experience in efficacy research in the field of physical education. The PI, via his faculty advisor, used a previously established email list of faculty members at PETE doctoral programs across the US (Boyce & Rickard, 2011). Experts were contacted via email and asked to complete a survey that reviewed and provided feedback based on the first draft of the instrument (see Appendix B). Of the 11 experts solicited for this review, 6 agreed to participate for a 54% response rate.

Evaluation criteria for experts. Experts were asked to evaluate items related to teaching task and provide feedback on their representation of the task constraints a physical educator may have when teaching a specific lifetime physical activities. Experts were also asked to review the instrument based on readability, clarity, conciseness, and overall layout. Expert review consisted of evaluation in three different areas: (a) participant perceptions of experience items, (b) teaching constraints influencing efficacy perceptions item, and (c) questions pertaining to clarity, readability, and conciseness of overall format (see Appendix C).

The first section included the two perception questions addressing teacher experience towards lifetime physical activities. Experts were asked to review each item and answer questions based on: (a) relevance to self-efficacy perceptions towards teaching of lifetime physical activities (*strongly disagree, disagree, neutral,* *agree, strongly agree*) (DeVellis, 2012), and (b) if the item was clear and concise (*strongly disagree, disagree, neutral, agree, strongly agree*). Experts were also given an open response category to address any content the construct failed to include. (see Figure 2).

The second section used the items extracted from the literature and the NASPE *Initial PETE Standards* (2008) addressing task constraints associated with teaching physical education activities. The teaching task constraints were listed and experts were asked to review each item based on: (a) relevance to the construct of self-efficacy perceptions towards teaching of lifetime physical activities (*strongly disagree, disagree, neutral, agree, strongly agree*) (DeVellis, 2012), and (b) if the item is clear and concise (*strongly disagree, disagree, neutral, agree, strongly agree*). Experts were given an open response question to address any content (task constraints) not addressed within the current items.

Scoring. Prior to the collection of responses from experts, a predetermined criterion was established for scoring. All items were scored on a 1-5 Likert scale, and quantitative questions with an average response score less than or equal to 3 were reviewed for clarification, elimination, or revision, and items that scored at least one low score (a one or two), were automatically reviewed. The PI used both the qualitative and quantitative feedback to refine the specified construct.

Revisions

Based on the recommendations of the experts, item content, definitions, and formatting were revised to make the survey an accurate representation of the construct, thus providing more evidence face and content validity. The final scale

11. How confident are you in your ability to <u>PRESENT (describe & demonstrate)</u> critical skill elements in the following activity categories?

	0- No Confidence	1	2	3	4	5- Moderately Confident	6	7	8	9	10 - Very Confident
Outdoor Pursuits (e.g., hiking, orienteering, rock climbing)	○	○	○	○	○	○	○	○	○	○	○
Fitness Activities (e.g., strength/weight training, yoga, cycling/biking)	○	○	○	○	○	○	○	○	○	○	○
Aquatic Activities (e.g., swimming, diving)	○	○	○	○	○	○	○	○	○	○	○
Individual Sports (e.g., track & field, self defense, gymnastics)	○	○	○	○	○	○	○	○	○	○	○
Dance/Rhythmic Activities (e.g., creative, line, modern)	○	○	○	○	○	○	○	○	○	○	○
Net/Wall Activities (e.g., tennis, badminton, pickleball)	○	○	○	○	○	○	○	○	○	○	○
Target Activities (e.g., golf, bowling, archery)	○	○	○	○	○	○	○	○	○	○	○

20. PEES-LPA question 11 is clear and concise.

☐ Strongly Disagree ☐ Disagree ☐ Neutral ☐ Agree ☐ Strongly Agree

21. **<u>Presentation</u>** of critical skill elements is an essential skill for instruction of lifetime physical activities.

☐ Strongly Disagree ☐ Disagree ☐ Neutral ☐ Agree ☐ Strongly Agree

22. Overall, this question is valuable in measuring the construct: *self-efficacy perceptions of PE teachers towards the instruction of lifetime physical activities.*

☐ Strongly Disagree ☐ Disagree ☐ Neutral ☐ Agree ☐ Strongly Agree

23. (Optional) Please provide any overall feedback you may have about question 11.

Figure 2. Sample item analysis questionnaire from expert review.

Table 7

Summary of Steps in Phase II

Step	Purpose	Source
1. Focus Group Review	• Test Pre-Pilot instrument for readability, clarity, format, and length. • Examine face & content validiy in the instrument's ability to represent the construct.	Bandura (2006) Hinkin (1995)
2. Expert Review	• Test Pre-Pilot instrument for readability, clarity, format, and length. • Examine face & content validiy in the instrument's ability to represent the construct.	Bandura (2006) DeVellis (2012)

consisted of 68 items with seven items added as a result of expert and focus group evaluation. The seven new questions represented a separation of the one adapted question into two distinct questions: (a) confidence to *differentiate* instruction in general physical education, and (b) confidence in *addressing the needs* of students with disabilities (see Table 7 for summation of Phase II).

Phase III: Validation Study

Phase III consisted of the following four steps: (a) instrument selection, (b) sampling framework, (c) sample size, and (d) recruitment procedures.

Instrument Selection

The instrument that was used in the Validation Study to measure self-efficacy perceptions was the *Physical Educator Efficacy Scale for Teaching Lifetime Physical Activities* (PEES-LPA). The PEES-LPA was comprised of: (a) demographic information, (b) perception items addressing teaching experience, skill ability, and psychological responses towards specific lifetime physical activities and (c) efficacy items addressing perceptions of explicit teaching behaviors during instruction of lifetime physical activities. The final scale consisted of 68 items (5 demographic, 21 perception of personal skill ability, and 42 self-efficacy to teach) (see Appendix D-PEES-LPA).

Sampling Framework

Groves, et al. (2009) state that how well a sample represents a population is highly dependent on the sampling frame, sample size, and recruitment procedures. The sampling framework for the Validation Study of the PEES-LPA consisted of four stages: (a) selecting the target population, (b) selecting the sampling frame population, (c) sampling of participants from sample frame, and (d) collection of responses from sample (Groves et al., 2013).

The target population is the ideal group of participants the PI wished to draw conclusion from (Fowler, 2009). The target population for the validation study was secondary physical educators within with membership association to either the national AAHPERD (now SHAPE America), or to their state AAHPERD.

Due to the fact that the target population is often theoretical and lacks specificity, and is often near impossible to obtain, a sampling frame is used to operationally define the target population for a study (Fowler, 2009; Groves et al., 2013). The sampling frame represents the individuals within the target population who have an *actual* chance of being selected for the study (Fowler, 2009). The sampling frame for the validation study consisted of the following criteria for participants: (a) were current in-service secondary physical education teachers (public or private), (b) were teaching with membership to either AAHPERD at the national or state level, (c) were on their representative states' local AAHPERD email listserve, and (d) agreed to volunteer to participate in the validation study.

The sample was the individuals selected from the sampling frame. This sample was the group of participants who were actually asked to participate in the validation study (Groves, et al. 2013). For the validation study of the PEES-LPA instrument, the sampling frame and sample went hand-in-hand. This was due to the fact that there is no comprehensive list of practicing secondary physical educators across the different state AAHPERDs.

The respondents who agreed to participate in the PEES-LPA validation study were 231 in-service secondary physical education teachers. A question was included prior to the online survey that asked participants to confirm that they are currently teaching secondary physical education. Those who selected “yes” were allowed to continue with the survey and remained in the sample. Those who selected “no” were denied access to the survey, and were thanked for their interest in participation. Due to this delimitation, there was a removal of 34 individuals, with another 15 removed due to incomplete data, for a total sample size of 182.

Sample Size

Evidence of construct validity of the PEES-LPA was determined based on results from an exploratory factor analysis. Using the relevant literature the PI gave careful consideration to the sample size of the validation study (Comrey, 1988; Field, 2009; Guadagnoli & Velicer, 1988; Tinsley & Tinsley, 1987). The PI concluded that for the validation study, a desired sample size of 200 participants would meet the acceptable standard, and would produce reliable results. The actual sample consisted of 182 in-service secondary physical education teachers.

Recruitment Procedures

The participants for the validation study were recruited to participate based on two primary methods: (a) receiving a solicitation email through participating state AAHPERD email listserve, or (b) in-person recruitment at the national AAHPERD convention in St. Louis, Missouri. The first step in the recruitment process was to contact the Executive Directors for all 50 state-level AAHPERD organizationn, to ask for their willingness and permission to distribute an email through their listserve requesting participants for the validation study. Permission to distribute the email, with embedded Qualitrics PEES-LPA link, was then granted by cooperating state Executive Directors.

The second form of recruitment took place at the 2014 National AAHPERD conference in St. Louis, Missouri. The PI conducted in-person solicitation to gather a list of email of in-service secondary physical educators agreeing to participate in the study, to then distribute the survey at a later time.

The emails that was sent out to physical educators, both through AAHPERD listserves and through in-person solicitation, made sure to specify that the survey was intended to measure efficacy beliefs of secondary physical education teachers (grades 6-12). The email that was sent out to physical educators and responses were collected in the Qualtrics online software (see Table 8 for a summation of the steps in Phase III).

Phase IV: Assessment of Reliability and Construct Validity

Because of the remaining uncertainty as to the number and nature of the factors underlying the items located in the PEES-LPA instrument, the current investigation remained exploratory in nature. Consequently, exploratory factor analysis (EFA) was conducted on items pertaining to physical education teacher's (N = 182) confidence to instruct lifetime physical activities. Exploratory factor analysis (EFA) has been found be valuable in the early stages of survey development due to its ability to describe and summarize the data according the factor correlations (Tabachnick, Fidell, & Osterlind, 2001).

Table 8

Summary of Steps in Phase III

Step	Purpose	Source
1. Instrument Selection	• To use the PEES-LPA to collect participant data on : demographic items, perception towards experience, and self-efficacy towards instruction of lifetime physical activities	
2. Sampling Framework	• Define target population, sampling frame, and sample	Groves et al. (2013) Fowler (2009)
3. Sample Size	• To collect responses from at least 200 participant; per EFA literature	Field (2000) Comrey (1988) Tinsley & Tinsley (1987) Costello & Osborne (2005)
4. Recruitment Procedures	• Elicit physical educators from the Southern District of AAHPERHD to participate in online survey •Contact state & county physical education director to help distribute PEES-LPA through their email system	AAHPERD (2013)

The following steps were used to evaluate assess the validity and reliability evidence for the PEES-LPA: (a) Step 1-test multivariate assumptions (missing values, univariate & multivariate outliers) (b) Step 2- test for sampling adequacy, (c) Step 3- choose factor extraction model, (d) Step 4- determine how many factors to retain (scree plot, Horn's [1965] parallel analysis, eigenvalues), (e) Step 5- factor rotation (varimx, oblim), (f) Step 6- examine factor structure, (g) Step 7- internal consistency test for reliability. All EFA analyses were entered and conducted in SPSS statistical software for analyzation.

Step 1- Test Multivariate Assumptions

Due to the fact that EFA with Principal Axis Factoring (the chosen extraction model) is sensitive to violations of multivariate assumptions, data were tested for: (a) missing data, (b) univariate normality, (c) univariate & multivariate outliers, and (d) size of item correlations (Tabachnick et al., 2001; Field, 2009).

Missing data. Missing data occurs in almost all behavior science research (Stevens, 2009). It is very common for educational researchers to use ad hoc methods (listwise/pairwise deletion) to deal with the problem (Peng, Harwell, Liou, & Ehman, 2006). Because the data from this EFA was found to be Missing at Random (MAR), and the PI seeking to retain all 182 participants, missing data were handled using an Expectation Maximization (EM) approach. The EM approach has two major phases:

> "In the expectation step, the process is similar to the regression-based imputation. First, starting values for the parameters (e.g., means, covariances) are obtained with available data. When this step

is completed, in the maximization step new values for the parameters are calculated with the newly imputed data along with the original observed data”(Schlomer, Bauman, & Card, 2010, p. 4).

The EM approach is seen in social science research provides an unbiased and efficient parameters, and is often see as particularly useful in EFA procedures and internal consistency procedures, both of which do not require hypothesis testing (Little, 1988).

Univariate normality. Univariate normality must be tested in the EFA process, due to the fact that it is a prerequisite for multivariate normality (Stevens, 2009). Univariate normality was examined by the visual examination of histograms for normal distribution, and the more formally through the evaluation of skewness (degree of symmetry of the distribution) and kurtosis (shape of the distribution against the normal distribution) statistics, and the Shapiro-Wilk’s test (detects wide variety of variation from normal distribution)(Stevens, 2009).

Univariate & multivariate outliers. Due to the fact that multivariate procedures are quite robust to violations of normality, it is important to examine the influence outliers may have on normality (Tabachnik et al., 2001). Univariate outliers were tested by examining z-scores and looking for extreme scores on each dependent variable (> 3.29, < -3.29). Any dependent variables found with extreme univariate outliers, were transformed using the logarithmic (log10) procedure.

Multivariate outliers were examined by the Mahalanobis distance statistic. The Mahalanobis distance demonstrates how far an individual participant response (case) is from the centroid of the all cases combined, regardless of sample size (Stevens, 2009). Any multivariate outliers found, were deleted.

Size of correlations. All variables were tested prior to running the EFA for extreme correlations. Visual inspection of the correlation matrix examined variables that loaded extremely high with other variables (>.85), or had correlations above .30 (some degree of correlation is needed for EFA) (Field, 2009; Stevens, 2009).

Step 2- Test for Sampling Adequacy

Sampling adequacy provides researchers information about how variables are grouped, thus better explaining the construct under investigation (Tabachnik et al., 2001). Measures of sampling adequacy evaluate whether items selected for a survey have some type of relationship to one another. Sampling adequacy for the PEES-LPA was tested through the Kaiser-Meyer-Olkin (KMO) measure, as well as testing the significance of Bartlett's test of sphericity, which tests if the correlation matrix produced is an identity matrix (Field, 2009; Stevens, 2009).

Step 3- Choose Factor Extraction Model

Principal axis factoring (PAF) is often noted for being the most widely used method of factor extraction when EFA is the chosen analytic method (Conway & Huffcutt, 2003). Pricipal axis factoring is a type of EFA that seeks to identify latent variables that contribute to the common variance among a set of selected variables, excluding variable-specific variance (Conway & Huffcutt, 2003; Field, 2009). Principal axis factoring "extracts the factors using a reduced correlation matrix, where the variance of each measure reflects its association with the other measures included in the factor analysis" (Russell, 2002, p. 1630). PAF was the chosen extraction model for this study.

Step 4- Determine How Many Factors to Retain

The resulting factor solutions were evaluated against the following criteria:

- Unrotated factors were required to satisfy Kaiser's (1958) criterion of eigenvalues being greater than 1.00
- Horn's (1965) parallel analysis
- Accepted configurations had to account for an acceptable percentage of total score variance
- Solutions to meet minimal Scree requirements (Field, 2009)
- Each factor rotated should include at least two appreciable factor loadings (i. e., ≥ .40)
- No more than 5% of the items should load on to more than one factor

Step 5- Rotate Factors

Rotations are used to help clarify and simplify the data structure (Stevens, 2009). Rotation does not improve the basic aspects of the results and factor structure, yet it influences the amount of variance extracted from each variable (Costello & Osborne, 2005). The two rotations used for the evaluation of the factor structure in the PEES-LPA were Varimax (orthogonal) and Oblim (oblique). The primary difference between the two rotations is that the Varimax rotation allows for the factors to remain uncorrelated, while the Oblim rotations keeps correlations among factors (Costellow & Osborne, 2005; Stevens, 2009). Both Varimax and Oblim rotations were considered for this study.

Step 6- Examine Factor Structure

After both rotations have been performed, resulting factor output should be examined for simple structure. Simple structure indicates that each variable in the survey helps to explain one, and only one, factor (area) of a construct (Costello & Osborne, 2005).

Step 7- Test Internal Consistency/Reliability

Internal consistency refers to the consistency of which the variables measure only one construct of interest, thus testing intercorrelation (Field, 2009). Internal consistency is widely used in EFA, and measured using Cronbach's alpha. Alpha is the "proportion of a scale's total variance that is attributed to a common source, presumably the true score of a latent variable underlying the items" (DeVellis, 2012, p. 37).

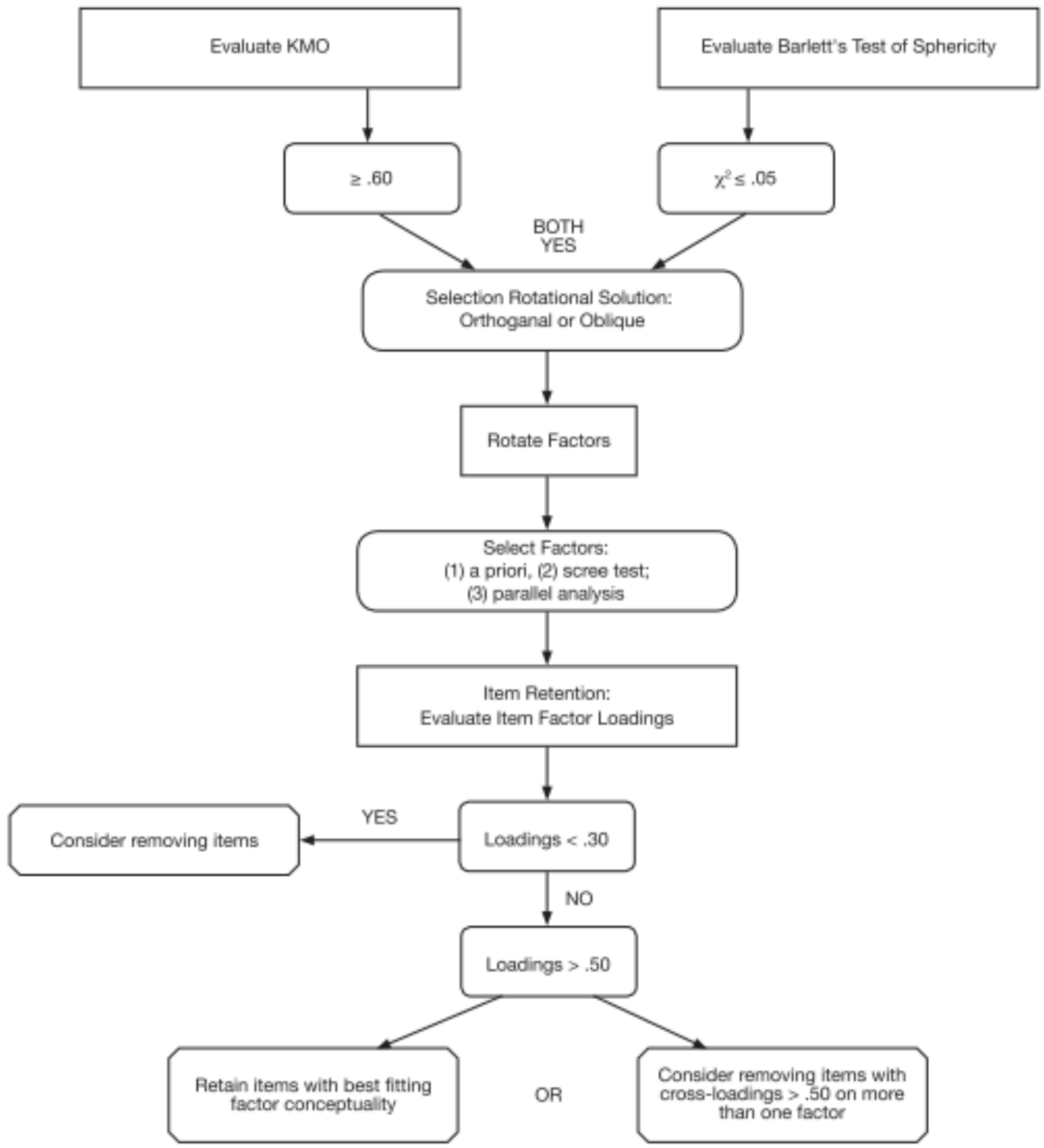


Figure 3. Flowchart of steps for conducting an exploratory factor analysis. Adapted from "Survey instrument validity part I: Principles of survey instrument development and validation in athletic training," L. J. Burton, & S. M. Mazerolle, 2011, *Education Research*, *6* (1), 27–35.

CHAPTER 4

RESULTS

The purpose of this study was to begin the process of validating a measure that could be used to predict physical educators' self-efficacy perceptions towards teaching lifetime physical activities. This instrument, the *Physical Educator Efficacy Scale for Teaching Lifetime Physical Activities* (PEES-LPA), was developed through expert review and pilot procedures. This study implemented both qualitative and quantitative procedures throughout the pilot and validation stages to address each of the research questions.

The two research questions (RQ) were addressed in the current study were:

RQ1: *Does the Physical Educator Efficacy Scale for Teaching of Lifetime Physical Activities (PEES-LPA) provide evidence for validity in measuring physical educator self-efficacy perceptions towards instruction of lifetime physical activities?*

RQ2: *Does the Physical Educator Efficacy Scale for Teaching of Lifetime Physical Activities (PEES-LPA) provide evidence for reliability in measuring physical educator self-efficacy perceptions towards instruction of lifetime physical activities?*

The preliminary pilot procedures used both quantitative and quantitative procedures to help determine the content and format of the newly created PEES-LPA instrument. The validation steps for the full instrument (PEES-LPA) utilized quantitative methodology to determine: (a) how many items to retain, (b) the factor structure of the latent variables, and (c) the reliability of internal consistency. Results will be presented in two major sections: pilot procedures and validation study. The results from the validation study will be further sub-divided sections that include: (a) descriptive statistics, (b) assumptions, (c) exploratory factor analysis, and (d) reliability.

Section I

Pilot Procedures

Focus Group

A focus group was conducted to qualitatively examine the instrument's content, readability, clarity, format, and length. A convenience sample of five individuals participated in the focus group, which consisted of 2 females and 3 males. The participants were doctoral students in physical education (adapted and pedagogy) from the University of Virginia, with diverse educational backgrounds representing the countries of the United States, South Korea, Ireland, and Jamaica.

The major points extracted from the focus group discussion that aided in instrument revision were:

- Revisions on 6 spelling errors
- The question referring to *adapting instruction* was changed to *differentiate instruction* due to the fact that it was referring to general physical education

classes, and the word *adapt* may prime participants towards individuals with disabilities.

- Questions were re-ordered based on the logical sequencing required to instruct a particular lifetime physical activity area (e.g., *identify* critical skill elements was moved up before *presenting* critical skill elements)
- The phrase "compared to the average physical educator" was included at the beginning of the question referring to ability to instruct a variety of activities within each activity area, to give participants a frame of reference to compare their own personal abilities.

Expert Review

Expert review consisted of distribution of the PEES-LPA in its draft form to both quantitatively and qualitatively evaluate items related to: (a) teaching task constraints a physical educator may have when teaching a specific lifetime physical activities, and (b) readability, clarity, conciseness, and overall layout of the items.

Eleven senior physical education pedagogy professors at doctoral-granting universities across the United States were requested to act as experts in reviewing the instrument. Of the eleven that were sampled, six agreed to participate, thus representing a response rate of 54%. Quantitative scores (on a 1-5 Likert scale) with responses ≤3 were reviewed for clarification, elimination, or revision. Additionally, items that scored at least one low score (a one or two) were automatically reviewed.

Quantitative analysis showed that four questions met the minimal standards for review and further examination:

1. Q8: The qualitative descriptors selected (novice, advanced beginner, competent, proficient, & expert) are appropriate for measuring personal skill ability? (*Towards lifetime physical activities*)
2. Q13: Is PEES-LPA Question 9 (competence to teach a variety of lifetime physical activities) clear and concise?
3. Q14: The qualitative descriptors selected (novice, advanced beginner, competent, proficient, & expert) are appropriate for measuring perception of teaching competence?
4. Q32: Is PEES-LPA item 14 (confidence in ability to adapt instruction in general physical education) clear and concise?

Due to expert review process being such a valuable part of the content and construct validity process (DeVellis, 2012), this PI chose to consider all qualitative feedback from experts. Selected qualitative expert feedback can be found it Table 9.

Based on the feedback from both the focus group and experts, item content, definitions, and formatting were revised to make the PEES-LPA more accurate representation of the construct, thus providing more evidence content and construct validity. The final scale consisted of 68 items (5 demographic, 21 perception of personal skill ability, & 42 self-efficacy to teach) with no items eliminated as a result of expert and focus group evaluation (see Table 9).

Table 9

Sample of Qualitative Feedback from Expert Review

Question for Experts	Responses
Q1: The activity categories (7) selected are clear and concise.	Expert #1: "The categories of target activities and net activities could have been combined into individual activities."
Q5: The format is appropriate for the selected demographic questions.	Expert #1: (*Regarding middle school being designated by the grades 5-8*) "In many cases, middle school doesn't begin until 6th or 7th grade. This might be a bit confusing for some teachers. Just a thought."
Q8: The qualitative descriptors selected (novice, advanced beginner, competent, proficient, & expert) are appropriate for measuring personal skill ability?	Expert#3: "I would add "no experience" category. This question might be problematic as some might have experience in a sport/activity but non in the others-have a hesitancy to teach or lack of competence to teach. Not sure how you get around that other than by listing individual sports."
Q14: The qualitative descriptors selected (novice, advanced beginner, competent, proficient, & expert) are appropriate for measuring perception of teaching competence compared to the average physical educator?	Expert #3: "Why is it important to compare to the average certified physical educator. I would just have them rate teaching competence." Expert #5: "What is the definition of an 'average' physical educator? I believe this word needs to be defined."
Q36-38: PEES-LPA item addresses the confidence in ability to meet needs of students with disabilities who are included in physical education classes (*for each of the lifetime physical activity areas*).	Expert #1: "I think assessment of students with disabilities is best done by an expert in adapted physical education, not necessarily the regular physical education teacher." Expert #2: "I would think that there would be some overlap between this question and the previous question (*adapting lessons for high/low skilled in general PE*)" Expert #4: "This one is a little tricky given the range of disabilities that are possible. It might be hard for the respondent to generalize" Expert #5: "Could this item be combined with item 14, as the ability to adapt a lesson should meet the needs of high and low skilled learners as well as students with disabilities."

Section II

Validation Study

Descriptive Statistics

A total of 231 participants agreed to participate in the survey. Of that sample, 34 individuals were removed due to being self-identified as being elementary physical educators (the Qualtrics online survey software collected no data for those 34 individuals). Additionally, 15 individuals were removed from the final sample due to incomplete surveys. The final sample of participants for this study comprised of 182 individuals (n=117, 64.3%, females; n=65, 35.7%, males).

Participants represented 24 states in the United States and one province in Canada, with the highest number of participants from Virginia (n=102), New York (n=12), Nevada (n=10), and South Carolina (n=9). Of these teachers, 41.2% (n=75) taught high school physical education, 44.5% (n=81) taught middle school, and 14.3% (n=26) taught at a combination of both high school and middle school. Participants had an average of 15.2 years of experience teaching physical education (SD=10.8, range=0 to 43), with 43.4% having 0-10 years of experience, 28% with 11-20 years of experience, 15.4% having 21-30 years of experience, 11.5% having 31-40 years of experience, and 1.6% having more than 40 years of experience. When the teachers were asked about their highest educational degree, 36.8% (n=67) indicated having a bachelor's degree, 62.1% (n=113) held a master's degree, and 1.1% (n=2) held a doctorate degree.

Assumptions

Prior to analysis, using the SPSS MAC 22.0 software data were screened in accordance with assumptions specific to factor analysis (Field, 2009) for missing values, both univariate and multivariate outliers, and for multivariate normality.

Missing values. A Missing Value Analysis (MVA) of the 68 variables showed that 20 of the variables reported 1 participant failing to answer an item (each representing 0.5% of data), and 4 variables had 2 participants failing to answer a survey item (each representing 1.1% of data). Examination of the missing data by participant showed that 24 of the 182 participants failed to self-report on one item of the survey. A follow-up Little's MCAR test (Little, 1988) resulted in a chi-square = 1471.99 (*df*=1479, *p*=.48), which indicates that the data were indeed missing at random (MAR), and no identifiable pattern exits to the missing data (Little, 1988).

Due to the missing data showing no identifiable patterns (MAR), exploratory factor analysis being the procedure to analyze the data, and this PI's attempt to avoid ad-hoc missing data procedures that reduce sample size (listwise & pairwise deletion), the Expectation-Maximization (EM) was used to handle the imputation of missing data based (Little, 1988; Peng, Harwell, Loiu, & Ehman, 2006). Upon completion of the EM procedure in SPSS, imputation of missing data was performed, thus yielding a complete dataset of 182 participants.

Univariate outliers. Univariate outliers were examined for all survey variables. Results revealed 30 potential outlier cases resulting from 18 variables (z's > 3.29, p < .001), with the following variables demonstrating more than one univariate outlier:

- Self-Efficacy to Identify Critical Skill Elements for Net/Wall Activities
- Self-Efficacy to Identify Critical Skill Elements for Target Sports
- Self-Efficacy to Assess Fitness, SE to Assess Individual Activities
- Self-Efficacy to Assess Net/Wall Activities
- Self-Efficacy to Give Feedback for Fitness
- Self-Efficacy to Give Feedback for Individual Activities
- Self-Efficacy to Give Feedback for Target Sports
- Self-Efficacy to Differentiate for Net/Wall Activities

Given the relatively large number of potential outliers in the dataset, evaluation of normality and possible transformation were considered prior to deletion of specific cases.

Univariate normality was examined separately for each variable through histograms, evaluation of skewness and kurtosis, and more formally through Shapiro-Wilk's test statistic. Values for the 45 variables *not* displaying univariate outliers were well within the acceptable limits for Skewness (<1.0) and Kurtosis (all values <1.2), and visual inspection of the histograms suggested approximate normal distribution. Results from the Shapiro-Wilk's test for these 45 variables did however indicate statistically significant departure from normality (p's < .05/63).

Of the 18 variables that displayed univariate outliers, 14 had moderate negative skew >1 (ranging from 1.0 to 1.64), and 10 variables had Kurtosis >1 (ranging from 1.0 to 3.87). Moreover, Shaprio-Wilk's tests of the 18 variables with univariate outliers revealed statistically significant departure from normality (p's < .05/63). Logarithmic (Log 10) transformations were applied to each of the 18 variables resulting in more normal distributions (skewness & kurtosis < 1.0). In addition, all follow-up univariate outlier analysis on these transformed variables failed to reveal any potential outliers (all Z's < 3.29) (see Appendix E). Due to the robustness exploratory factor analysis has towards violations of normality, especially when no observed outliers influence said normality, as well as the normal distribution found in skewness and kurtosis (<1.0), the variables were left unaltered (Gorsuch, 1983; Stevens, 2009).

Multivariate outliers. Multivariate outlier analyses were conducted separately for the two groups on the combined dependent variables. Mahalanobis Distance (MD) initially revealed 10 multivariate outlier cases compared to the chi-squared critical value of 104.7 (df = 63, α = .001). Due to fact that multivariate outliers can hide behind other multivariate outliers (Nunnally, 1978; Stevens, 2009), one-by-one each case with a MD > 104.7 was deleted and the new model (post-deletion) was re-evaluated for multivariate outliers. Overall, 13 multivariate outliers were detected, and then deleted from the overall sample.

A new dataset, with 13 deleted multivariate outliers (N=169), was saved and compared to the full model (13 outlier cases included, N=182) using exploratory factor analysis (principal axis factoring) with both orthogonal and oblique rotations. A side-by-side comparison of the two models (with and without deletion of multivariate outliers) revealed that multivariate outliers had no effect on the factor structure or factor loadings (see Table 10). Due to 13 multivariate outliers having no effect on the factor structure and loadings, paired with the univariate normality

Table 10

Comparison of EFA Models: With and Without Inclusion of Multivariate Outliers (PAF, Varimax Rotation)

Factor	1 (Full Model)	1 (MV outliers removed)	2 (Full Model)	2 (MV outliers removed)	3 (Full Mode)l	3 (MV outliers removed)	4 (Full Model)	4 (MV outliers removed)	5 (Full Model)	5 (MV outliers removed)	6 (Full Model)	6 (MV outliers removed)
Tran_Assess_Target	0.836	0.837										
Tran_FB_Target	0.831	0.832										
Tran_FB_Net	0.818	0.825										
Tran_Assess_Net	0.816	0.823										
Trans_Present_Net	0.803	0.81										
Present_Target	-0.783	-0.808										
Tran_Identify_Net	0.773	0.801										
Tran_Identify_Target	0.77	0.776										
Trans_Diff_Target	0.768	0.771										
Tran_Diff_Net	0.724	0.716										
Variety_Target	-0.628	-0.627										
Tran_Variety_Net	0.62	0.624										
SkillAbilty_Target	-0.609	-0.597										
Tran_SSN_Target	0.586	0.575										
SkilAbility_Net	-0.58	-0.562										
Tran_SSN_Net	0.565	0.555										
Freq_Target	-0.484	-0.531										
Freq_Net												
Identify_Dance			0.931	0.948								
Present_Dance			0.928	0.934								
Assess_Dance			0.925	0.931								
Feedback_Dace			0.917	0.93								
Diff_Dance			0.896	0.89								
Variety_Dance			0.836	0.846								
SkillAbility_Dance			0.77	0.779								
SSN_Dance			0.743	0.723								
Freq_Dance			0.725	0.722								
Present_Aquatic					0.908	0.909						
Identify_Aquatic					0.869	0.863						
Assess_Aquatic					0.856	0.859						
Feedback_Aquatic					0.846	0.857						
Variety_Aquatic					0.84	0.855						
SkillAbility_Aquatic					0.829	0.834						
Diff_Aquatic					[illegible]	[illegible]						

Factor	1 (Full Model)	1 (MV outliers removed)	2 (Full Model)	2 (MV outliers removed)	3 (Full Mode)l	3 (MV outliers removed)	4 (Full Model)	4 (MV outliers removed)	5 (Full Model)	5 (MV outliers removed)	6 (Full Model)	6 (MV outliers removed)
Present_Outdoor							0.897	0.917				
Feedback_Outdoor							0.895	0.913				
Assess_Outdoor							0.886	0.908				
Diff_Outdoor							0.857	0.869				
Identify_Outdoor							0.838	0.86				
Variety_Outdoor							0.824	0.808				
SkillAbility_Outdoor							0.765	0.751				
SSN_Outdoor							0.719	0.735				
Freq_Outdooor							0.572	0.575				
Tran_Identify_Fit									0.834	0.824		
Tran_Present_Fitness									0.815	0.815		
Tran_Assess_Fitness									0.795	0.803		
Trans_Diff_Fitness									0.791	0.803		
Tran_FB_Fitness									0.782	0.764		
Tran_Variety_Fitness									0.689	0.717		
SkillAbility_Fitness									-0.664	-0.705		
Tran_SSN_Fit									0.655	0.575		
Trans_Freq_Fit									0.44	0.464		
Identify_Ind											0.804	0.79
Tran_Assess_Ind											-0.766	-0.769
Variety_Ind											0.761	0.768
Tran_FB_Ind											-0.757	-0.762
Present_Ind											0.748	-0.744
SkillAbility_Ind											0.677	0.737
Trans_Diff_Ind											-0.669	0.687
SSN_Ind											0.641	0.634
Freq_Ind											0.479	0.426

found within skewness, kurtosis, and z-scores, this PI chose to retain the full model (*N*= 182) with inclusion of the multivariate outliers.

Exploratory Factor Analysis

Principal axis factor (PAF) extraction was conducted on all 63 variables, and both varimax and oblim rotations were considered in attempt to uncover simple structure. PAF was chosen as the extraction method due to a slight deviation in the data from normality. The number of factor to be retained was based on a fulfilling a variety of considerations, including: (a) minimal factor retention score of .40, (b) parallel analysis, (c) Kaiser's Eignevalue rule (greater than 1), (d) examination of the scree plot, (e) the amount of variance accounted for by the solution that are retained (> 50%), (f) no more than 5% of the items should load on to more thank one factor, and (g) results should have good internal consistency reliability and interpretability.

Prior to performing the PAF, the dataset was screened to ensure accuracy of the data, and to verify its suitability for factor analysis. The Kaiser-Meyer-Olkin (KMO) test for sampling adequacy (KMO = .890) was found meet the .60 minimal standard (Field, 2009), as well as Bartlett's test of sphericity rejected null hypothesis that the correlation matrix is an identity matrix, therefore data meets minimal standard ($\chi 2$= 15413.4, df = 1953, $p < .001$). Additionally, variable correlations were examined for extreme correlations and for enough correlations to warrant factor analysis ($r > 0.3$) (Field, 2009) (see Appendix F- Correlation Matrix) A visual inspection of the correlation matrix revealed minimal low extreme correlations, and the majority of correlations meeting the minimal benchmark for factor analysis ($r > 0.3$). A visual inspection of the high-end extremes showed that the *Identify* and *Present* variables, as well as the *Feedback* and *Assess* variables did have extremes that possibly influenced multicollinearity (>.85). Due to the exploratory nature of the research design the items were left in the sample.

Principal Axis Factor extraction was performed using SPSS MAC 22.0 on the 63 items in the PEES-LPA. As a preliminary step, principal component extraction revealed the presence of 10 factors with eigenvalues greater than 1.0, which accounted for 81.7% of the total observed score variance in the unroated model. Examination of the resulting structure matrix failed to reveal a clear pattern of simple structure across the 10 factors (see Appendix G- Unrotated Factor Matrix).

A visual inspection of the scree plot revealed a clear point of inflexion after the sixth factor, supporting the retention of only 6 factors (see Figure 4- Scree Plot). This was in comparison to a follow-up Horn's Parallel Analysis, which suggested the need to retain only 8 factors in the final model. As a result, both varimax (orthogonal) and oblim (non-orthogonal) rotations were examined for simple structure using both 6 and 8 extracted factors.

The unrotated PAF extraction results indicated that the eight-factor solution accounted for 74.1% of the total observed score variance (see Appendix H). Again, no clear simple structure pattern was revealed. As a result, a Varimax rotation was conducted on the 8-factor solution. The first six factors accounted for an appreciable amount of the observed score variance 14.5%, 11.8%, 11.2%, 10.5%, 9.3%, and 8.7% respectively, whereas the final the final two factors (7 & 8) only explained 5.1% and 2.7% of the variance. An examination of the structure matrix of the eight-factor solution revealed simple structure for the first six factors, and

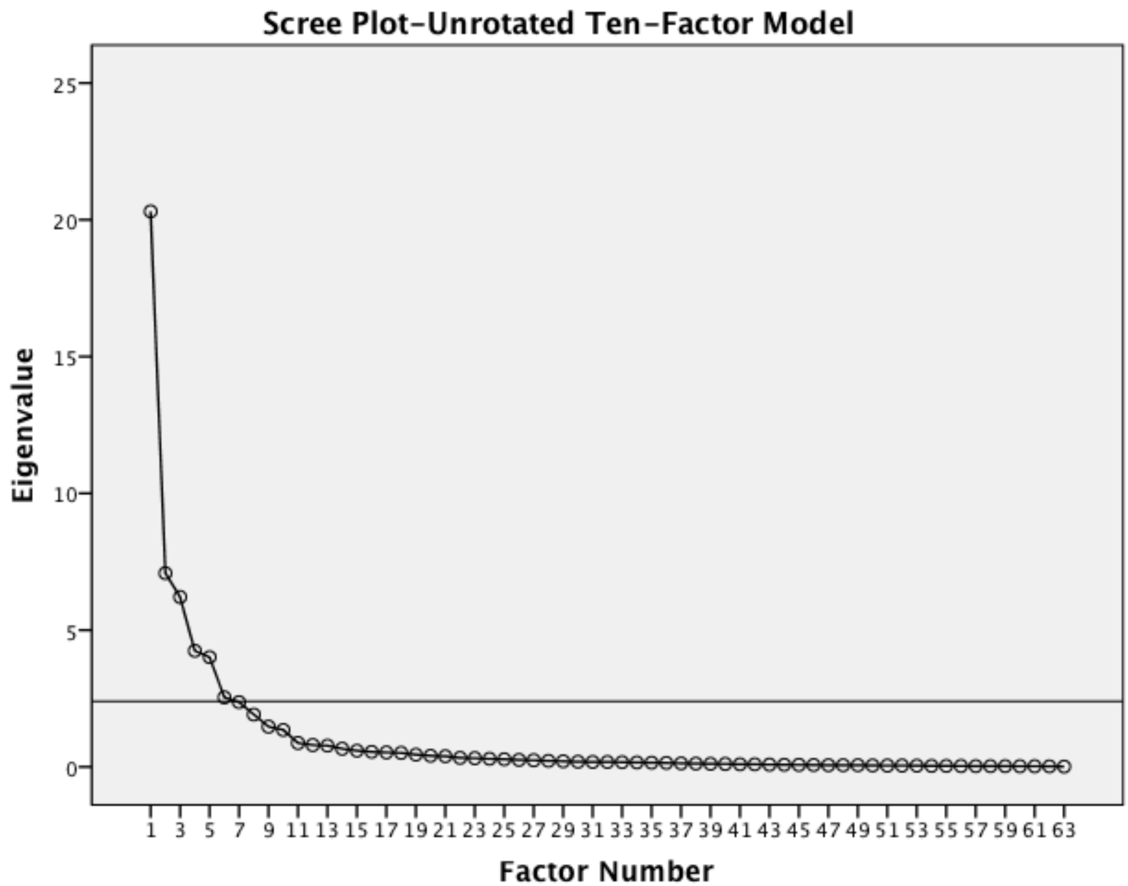


Figure 4. Scree plot of ten-factor unrotated model.

Table 11

Varimax Rotated Factor Matrix: Eight-Factor PAF[a]

	Factor							
	1	2	3	4	5	6	7	8
Tran_FB_Target	.867							
Tran_Assess_Target	.852							
Present_Target	-.826							
Tran_Identify_Target	.814							
Trans_Present_Net	.777							
Tran_Assess_Net	.774							
Tran_FB_Net	.771							
Trans_Diff_Target	.736						.412	
Tran_Identify_Net	.732							
Variety_Target	-.649							
Tran_Diff_Net	.637						.501	
SkillAbilty_Target	-.625							
Tran_Variety_Net	.588							-.548
SkilAbility_Net	-.554							.516
Freq_Target	-.465							
Identify_Dance		.932						
Present_Dance		.927						
Assess_Dance		.921						
Feedback_Dace		.917						
Diff_Dance		.885						
Variety_Dance		.858						
SkillAbility_Dance		.784						
SSN_Dance		.736					-.418	
Freq_Dance		.733						
Present_Aquatic			.915					
Identify_Aquatic			.876					
Assess_Aquatic			.865					
Feedback_Aquatic			.855					
Variety_Aquatic			.841					
Diff_Aquatic			.828					
SkillAbility_Aquatic			.827					
SSN_Aquatic			.748					
Freq_Aquatic			.491					
Present_Outdoor				.902				
Feedback_Outdoor				.896				
Assess_Outdoor				.887				
Diff_Outdoor				.848				
Identify_Outdoor				.841				
Variety_Outdoor				.838				

Varimax Rotated Factor Matrix: Eight-Factor PAF[a]

	Factor							
	1	2	3	4	5	6	7	8
SkillAbility_Outdoor				.774				
SSN_Outdoor				.713			-.459	
Freq_Outdooor				.571				
Tran_Identify_Fit					.833			
Tran_Present_Fitness					.827			
Tran_Assess_Fitness					.793			
Tran_FB_Fitness					.785			
Trans_Diff_Fitness					.755			
Tran_Variety_Fitness					.730			
SkillAbility_Fitness					-.687			
Tran_SSN_Fit					.599		.559	
Trans_Freq_Fit					.437			
Variety_Ind						.804		
Identify_Ind						.801		
Tran_Assess_Ind						-.761		
Tran_FB_Ind						-.753		
Present_Ind						.747		
SkillAbility_Ind						.695		
Trans_Diff_Ind						-.622	.388	
SSN_Ind						.583	-.530	
Freq_Ind						.488		
Tran_SSN_Net	.438						.736	
Tran_SSN_Target	.497						.654	
Freq_Net								.438

Extraction Method: Principal Axis Factoring.
Rotation Method: Varimax with Kaiser Normalization.[a]

a. Rotation converged in 7 iterations.

Table 12

Varimax Rotated Factor Matrix: Six-Factor PAF[a]

	Factor					
	1	2	3	4	5	6
Tran_Assess_Target	.836					
Tran_FB_Target	.831					
Tran_FB_Net	.818					
Tran_Assess_Net	.816					
Trans_Present_Net	.803					
Present_Target	-.783					
Tran_Identify_Net	.773					
Tran_Identify_Target	.770					
Trans_Diff_Target	.768					
Tran_Diff_Net	.724					
Variety_Target	-.628					
Tran_Variety_Net	.620					
SkillAbilty_Target	-.609					
Tran_SSN_Target	.586					
SkilAbility_Net	-.580					
Tran_SSN_Net	.565					
Freq_Target	-.484					
Freq_Net	-.278					
Identify_Dance		.931				
Present_Dance		.928				
Assess_Dance		.925				
Feedback_Dace		.917				
Diff_Dance		.896				
Variety_Dance		.836				
SkillAbility_Dance		.770				
SSN_Dance		.743				
Freq_Dance		.725				
Present_Aquatic			.908			
Identify_Aquatic			.869			
Assess_Aquatic			.856			
Feedback_Aquatic			.846			
Variety_Aquatic			.840			
SkillAbility_Aquatic			.829			
Diff_Aquatic			.820			
SSN_Aquatic			.735			
Freq_Aquatic			.495			
Present_Outdoor				.897		
Feedback_Outdoor				.895		
Assess_Outdoor				.886		

Varimax Rotated Factor Matrix: Six-Factor PAF[a]

	Factor					
	1	2	3	4	5	6
Diff_Outdoor				.857		
Identify_Outdoor				.838		
Variety_Outdoor				.824		
SkillAbility_Outdoor				.765		
SSN_Outdoor				.719		
Freq_Outdooor				.572		
Tran_Identify_Fit					.834	
Tran_Present_Fitness					.815	
Tran_Assess_Fitness					.795	
Trans_Diff_Fitness					.791	
Tran_FB_Fitness					.782	
Tran_Variety_Fitness					.689	
SkillAbility_Fitness					-.664	
Tran_SSN_Fit					.655	
Trans_Freq_Fit					.440	
Identify_Ind						.804
Tran_Assess_Ind						-.766
Variety_Ind						.761
Tran_FB_Ind						-.757
Present_Ind						.748
SkillAbility_Ind						.677
Trans_Diff_Ind						-.669
SSN_Ind						.641
Freq_Ind						.479

Extraction Method: Principal Axis Factoring.
Rotation Method: Varimax with Kaiser Normalization.[a]

a. Rotation converged in 6 iterations.

double loadings for the seventh and eighth, when loadings ≥ 0.40 were considered (see Table 11). As a result of the evidence, the seventh and eighth factors were then removed, thus supporting the use of six-factor model.

An examination of the unrotated six-factor solution indicated that 61.8% of the total observed score variance was accounted for in the model, though the resulting structure matrix failed to reveal simple structure across the six factors. Varimax rotation indicated that the six factors accounted for 67.9% of the cumulative score variance, and appreciable amounts of variance for each of the six factors: 15.5%, 11.8%, 10.9%, 10.5%, 9.7%, 9.2%. An examination of the structure matrix of the six-factor solution revealed simple structure for the when loadings ≥ 0.40 were considered (see Table 12).

A final analysis was conducted using an oblim rotation. Results revealed that cumulatively, the model accounted for 67.8% of the total subset score variance. Additionally, results revealed a similar pattern of variable loadings across the six factors as indicated in the pattern matrix of coefficients (See Appendix I). Due to the simple structure being found in the varimax (orthogonal) rotation, this model was selected to use in results interpretation.

The final model revealed one factor (factor 1) was defined by 18 subsets, while the other five factors (factors 2-6) were defined by 9 subsets with appreciable loadings ≥.0.40. The breakdown of item factor loadings were:

- Factor 1-Items specific to Target Sports & Net/Wall Activities (N= 18)
- Factor 2-Items specific to Dance & Rhythmic Activities (N=9)
- Factor 3- Items specific to Aquatic Activities (N=9)

- Factor 4- Items specific to Outdoor Pursuits (N=9)
- Factor 5- Items specific to Fitness Activities (N=9)
- Factor 6- Items specific to Individual Sport Activities (N=9)

  (see Table 52)

The factor correlation matrix revealed a moderately large correlation (i.e., greater then .32, or 10% overlap of variance in factors) (Tabachnick, et al., 2001*)* for factors 1 and 4 (.475), 1 and 5 (.356), 1 and 6 (-.435), and 3 and 4 (.328). Correlations between the remaining factors were notably smaller (all < .27) (see Table 13).

Reliability

Cronbach's alpha was used to measure internal consistency of the 63-item self-efficacy instrument. Reliability results for the full instrument was Cronbach's α=.95. Internal consistency was also measured for each of the six factors (subscales) with results, which also resulted in high alpha scores overall:

- Factor 1-Items specific to Target & Net/Wall Activities (*N*=18): α=.949
- Factor 2-Items specific to Dance & Rhythmic Activities (*N*=9): α=.955
- Factor 3- Items specific to Aquatic Activities (*N*=9): α=.953
- Factor 4- Items specific to Outdoor Pursuits (*N*=9): α=.944
- Factor 5- Items specific to Fitness Activities (*N*=9): α=.924
- Factor 6- Items specific to Individual Sport Activities (*N*=9): α=.941

Due to the high internal consistency alpha, and the potential for multicollinearity influencing reliability estimates, 10 follow-up internal consistency examinations were done to test the reliability estimates of the full model when

Table 13

Factor Correlation Matrix: Six Factor Model

Factor	1	2	3	4	5	6
1	1.000	-.270	.255	.475	.356	-.435
2	-.270	1.000	-.158	-.196	-.119	.262
3	.255	-.158	1.000	.328	.064	-.244
4	.475	-.196	.328	1.000	.258	-.242
5	.356	-.119	.064	.258	1.000	-.238
6	-.435	.262	-.244	-.242	-.238	1.000

Extraction Method: Principal Axis Factoring.
Rotation Method: Oblimin with Kaiser Normalization.

variables were removed. Removal of variable groupings provided the following reliability estimate results:

- Full model with the exclusion of *Identify* (*N*= 7) variables: α=.955
- Full model with the exclusion of *Present* (*N*= 7) variables: α=.955
- Full model with the exclusion of *Assess* (*N*= 7) variables: α=.954
- Full model with the exclusion of *Feedback* (*N*= 7) variables: α=.953
- Full model with the exclusion of *Identify* (*N*= 7) & *Assess* (*N*= 7) variables: α=.947
- Full model with the exclusion of *Identify* (*N*= 7) & *Feedback* (*N*= 7) variables: α=.946
- Full model with the exclusion of *Present* (*N*= 7) & *Assess* (*N*= 7) variables: α=.947
- Full model with the exclusion of *Present* (*N*= 7) & *Feedback* (*N*= 7) variables: α=.946
- Full model with the exclusion of *Net/Wall* (*N*= 9) variables: α=.956
- Full model with the exclusion of *Target* (*N*= 9) variables: α=.955

A brief examination of the self-efficacy results (see Appendix K) revealed: (a) perceptions of skill-ability were highest in fitness, net/wall, and target activities, (b) frequency of participation over the last year for all seven activity area showed very similar means (ranging from 2.3-2.8, on a 0-5 scale), (c) physical educators tended to be least confident when instructing students with special needs, and (d) almost unanimously for each of the six self-efficacy questions, participants were most confident in instruction of fitness, target, and net/wall activities.

CHAPTER 5

DISCUSSION

The purpose of the present study was to provide validation for, and explore the factor structure of, the Physical Educator Efficacy Scale for Teaching Lifetime Physical Activities (PEES-LPA). In order to satisfy this objective, data were collected from 182 in-service secondary physical education teachers who agreed to participate in the survey. The resulting instrument, PEES-LPA, is a 68-item survey composed of (a) demographic questions, (b) items measuring perception towards personal ability, and (c) items measuring self-efficacy towards instruction of lifetime physical activities. An exploratory factor analysis was used in order to explore the appropriateness of scale item placement, factor structure, and internal consistency.

The major findings from this study demonstrated that: (a) the resulting factors exhibited simple structure that aligns with literature supporting the classification of lifetime physical activities (AAHPERD, 2013), (b) factors were composed of items that logically relate, as well as showed high levels of internal consistency, and (c) preliminary results showed that the PEES-LPA appears to be an appropriate instrument for measuring self-efficacy perceptions of physical educators, though item reduction (while still upholding high internal consistency) may be a logical approach to future research. While these overall findings are positive, they should be considered cautiously with the nature of the sample. This chapter will include: (a) a discussion of the results of this study, (b) recommendations for future research, and (c) conclusion.

Discussion

The results from this study show both the quantitative and qualitative methods used to employ validation and reliability procedures on the PEES-LPA instrument. Overall, 231 individuals agreed to participate in the full PEES-LPA survey, though only 182 of those participant responses had sufficient data to include in the final sample. The majority of those participants that were excluded from the final sample self-reported to only teach elementary physical education. Due to the uneven breakdown of males and females and states represented in the sample, paired with the study's research design that may have exhibited bias towards participants who would already perceive themselves to be confident and professionally engaging, a larger and more representative sample (e.g., 400+) would be desired.

Of the participants sampled, over 90% were solicited through state AAHPERD email listserve, compared to the 8-10% that came from in-person solicitation. The online mass distribution demonstrated to be a more expedient method for distribution, especially due to teachers only having to click on one link in an email. It is important to note that this third-party email distribution acted as a limiting factor due to each participating state AAHPERD only sending one blanket email to its members, with no subsequent reminder emails. Additionally, teachers from state of Virginia represented 56% of the total participant sample, a direct reflection of the state this study was administered, with the next highest state, New York, only representing 6.5%.

Validation of Survey Instrument

RQ1: Does the Physical Educator Efficacy Scale for Teaching of Lifetime Physical Activities (PEES-LPA) provide evidence for validity in measuring physical educator self-efficacy perceptions towards instruction of lifetime physical activities?

The PEES-LPA was the first documented attempt to construct an instrument specific to secondary physical educator self-efficacy perceptions towards the instruction of lifetime physical activities. The PEES-LPA was constructed and validated using the recommendations and guidelines from the provided instrument development literature, paired with Bandura's *Guide for Constructing Self-Efficacy Scales* (2006) and DeVillis's *Scale Development: Theory and Practice* (2012).

In survey development, validity is almost universally seen as the primary standard for assessing the adequacy of a group of variables (Johnson, 2011). Though its counterpart reliability can often easily be tested through established statistical techniques, the understanding of validity is often more subjective and complex (Johnson, 2011; MackKenzie, et al., 2011). The four major sources of validity most often associated with survey development are: (a) face, (b) content, (c) criterion, and (d) construct (DeVellis, 2012). The PEES-LPA was assessed using pilot and validation procedures to help provide evidence for face, content, and construct validity.

Due to the fact that face validity is a sub-division of content validity (DeVellis 2012; Lucko & Rojas, 2009), the two will be addressed in this discussion simultaneously. Additionally, no criterion-related validation procedures were conducted in testing of the PEES-LPA instrument due to the exploratory nature of the research design, and lack of any other instrument close enough in nature to align for predictive or concurrent validity (DeVellis, 2012; Hinkin, 1995; Litwin 1995).

Content & face validity of the PEES-LPA. In survey development research, face and content validity are most often qualitative measures of the relevancy and representativeness of an instrument's variables compared to the construct of interest (Haynes, et al., 1995; Messick 1995). Content validity evidence is also understood by many researchers to be the prerequisite for establishing construct validity (DeVellis, 2012; Litwin, 1995).

Content validity evidence for the PEES-LPA came from conducting a focus group and through an evaluative expert review process. The focus group consisted of University of Virginia doctoral students in physical education ($N = 5$) (pedagogy and adapted), and expert reviewers ($N = 6$) were physical education pedagogy professors at doctoral-granting institutions from across the United States. Both groups were asked to evaluate survey items related to teaching tasks selected, lifetime physical activity areas, as well as for readability, clarity, and conciseness.

Results from focus group were beneficial in addressing many preliminary problems such as wording, question order, and grammatical errors. These findings are synonymous with literature addressing focus groups feedback in the early stages of scale development (DeVellis, 2012). In comparison, the expert review feedback was centered on "big-picture" theoretical questions addressing the defining characteristics of the PEES-LPA construct. These trends are directly in line with much of the literature supporting the impact of focus groups and expert review in survey development (Bandura, 2006; DeVellis, 2012; Lucko & Rojas, 2010).

One interesting finding from the expert review process was that when experts were given the specific quantitative parameters (e.g., 1-5 scale) to frame their evaluation, they were all generally in agreement as to each of the questions on the PEES-LPA. Though when this PI evaluated the open-ended responses from each expert, there was wide variation in how they felt the PEES-LPA items could be improved to help align more closely with the selected construct. As a result of this disagreement, a follow-up technique known as cognitive testing, where potential respondents are asked to evaluate the questions empirically, may have helped clarify discrepancy (Bandura, 2006; Groves et al., 2013; DeVellis, 2012). An additional problem that arose from this process was that way the expert evaluation form failed to provide information about question strength (i.e., the ones the participants can answer most accurately) (Groves, et al., 2013). To address this issue, an additional question should have been added asking them to rank the questions on relevance.

Construct validity. Construct validity in survey development research is concerned with the theoretical relationship of a variable (e.g., item on a scale) to some other variable (DeVellis, 2012). Cronbach and Meehl (1955) as well as Clark and Watson (1995), both emphasize that the best, and most efficient, measures are those with established construct validity; "they are manifestations of constructs in an articulated theory that is well supported by empirical data. Thus, construct validity lies at the heart of the clinical utility of assessment and should be respected by scale developers and users alike" (Clark & Watson, 1995, p. 310). Evidence of construct validity in the validation of the PEES-LPA came from the review processes in the pilot procedures, as well as the statistical examination of the factor structure using exploratory factor analysis.

Testing of EFA assumptions. This section will discuss results associated with the assumptions tested prior to the exploratory factor analysis, and will be broken down in to three major sections: (a) sample size, (b) missing values, and (c) univariate & multivariate outliers.

Sample size. When examining the literature specific to sample size requirements for exploratory factor analysis, the only consensus that resonates is that a sample should be "large". Therefore, the larger the sample size, the greater likelihood of increased normality probability, as well as overall statistical significance (DeVellis, 2012; Hinkin, 1995). What specifically does "large" mean? Unfortunately, this question can be complicated for researchers to come to a consensus on. The issue here lies in fact that the lack of agreement has confused researchers, and has even started to create a divided opinion between what is *acceptable*, and what is *practical* (Costello & Osborne, 2005, DeVellis, 2012).

The final sample for this study consisted of 182 participants with 63 variables used in the factor analysis, for a final item-to-participant ratio of 1/2.9. In many researching circles this sample size would fall below acceptable limits. For example, Tinsley & Tinsley (1987) recommended 5-10 participants per variable, Field (2009) recommend 10 participants per variable, and Comrey (1988) recommend 200 participants.

What these recommendations fail to take in to account is the nature of the data, and the complex dynamics of factor analysis (Costello & Osborne, 2005).
"In general, the stronger the data, the smaller the sample can be for an accurate analysis. Strong data in factor analysis means uniformly high communalities without

crossloadings, plus several variables loading strongly on each factor" (Costello & Osborne, 2005, p. 4). Gualdagnoli & Velicer (1988) further supported this point by illustrating that when communalities are quite high (> .60), and correlation coefficients are > .80, then smaller sample sizes are acceptable.

An evaluation of the factor loadings, and communalities extracted from the EFA of the PEES-LPA it is clear that simple structure was evident from the early stages in the process. Furthermore, both the final models (Varimax & Oblim) when six-factors were retained, showed (1) no crossloadings, and (2) very high factor loadings (most loadings >.50) for each factor. Though prior to participant recruitment the desired sample size was delimited to 200+, as a result of the factor structure, correlations, and factor scores, the 182 participants are deemed acceptable.

Missing values. During the early stages of the EFA process, it was the objective of this PI to retain as many participants as possible. The rationale for this reasoning stemmed from: (a) the low final sample size, (b) the high number of dependent variables, and (c) attempting to avoid statistical software that uses default ad-hoc deletions to eliminate any participant who fails to answer one question.

This study employed an Exepectation Maximzation (EM) method to handle the imputation of missing data. The EM method used a series of regressions in SPSS for to compute an estimate from the patterns in the data for each missing data. This method has been widely used in educational research, and is even preferred by some researchers over ad-hoc deletions (Peng, et al., 2006). Using the EM method, the final sample size consisted of 182 participants, compared to an ad-hoc procedure, such as a listwise deletion, which would have retained a final sample size 158.

Univariate & Multivariate outliers. Upon initial examination of univariate outliers in the dataset, 18 outliers were found (z's > 3.29). Multivariate procedures are particularly sensitive to univariate outliers, so variable transformations were conducted. An examination of the historgrams and skewnewss statistics revealed that all 18 variables showed moderate negative skew, showing participants overall rated their self-efficacy levels quite high for those variables. The breakdown of the variables transformed was: Fitness variables (*N*=6), Net/Wall variables (*N*= 6), Target sport variables (*N*= 4), and Individual performance variables (*N*=2). Though this analysis will not specifically examine the self-efficacy results from the teachers, it worth noting that these variables were skewed towards higher scores (e.g., teachers rating themselves high in self-efficacy towards fitness).

Tabachnick et al., (2001), discussed that data transformation is a common practice, and should be seen more as data re-expression, rather than transformation. The one issue that did present itself in the results was that the Shapiro-Wilk's test for normality for *all* 63 variables did show significant departure from normality (p's < .05/63). For most multivariate research, this would be a red flag that would significantly hinder proceeding with future analyses (Field, 2009; Hinkin, 1995). Due to the (a) robustness exploratory factor analysis has towards violations of normality, (b) no observed outliers influence said normality, as well as (c) the normal distribution found in skewness and kurtosis (<1.0), the variables were left unaltered (Gorsuch, 1983; Stevens, 2009)

The process used to handle the multivariate outliers was not quite as simplistic as the transformations for univariate outliers. The results showed that overall, there were 13 total multivariate outlier cases in the dataset (Mahalanobis distance (MD) = 104.7, p=.0001). Mahalanobis distance is also highly sensitive to outliers, and single extreme observations, or groups, departing from the main data structure can have severe effect on MD (Hinkin, 1995). As a result, starting with the individual case with the highest MD, one-by-one each multivariate outlier was deleted. After the deletion of each outlier, the model was retested for any outlier that may have hidden behind another (Hinkin, 1995).

Due to the sensitive sample size, eliminating 13 cases was a cause for concern. As a result, the PI chose to do a side-by-side analysis of the full model (N= 182), with the new model after deletion of the 13 multivariate outliers (N=169). An examination of the new model showed the nearly the exact same simple structure, factor scores, and reliability estimates. Due to 13 multivariate outliers having minimal effect on the factor structure and loadings, paired with the univariate normality found within skewness, kurtosis, and z-scores, this PI chose to retain the full model (N= 182) with inclusion of the multivariate outliers.

Exploratory factor analysis. Principal axis factoring (PAF) extraction was performed, and both varimax and oblim rotations were considered in an attempt to uncover simple structure. Both theoretical and empirical evidence was considered when deciding the model and number of factors to retain.

Model selection. As with most EFA research, the unrotated full model rarely produces simple structure (Field, 2009). Thus, Horn's (1965) parallel analysis, widely considered to be the gold standard for determining factor structure, was examined (Conway & Huffcutt, 2003). Horn's parallel analysis reveled that the ideal model should retain eight factors. Examination of both Varimax and Oblim rotations with a delimited eight-factor model brought about factor structure and loadings that needed further examination.

First, the seventh and eighth factors produced double loadings for 12 variables, each of which had higher factor loadings in factors 1-6. Secondly, visual inspection showed that there was apparent simple structure for the first six factors, with the seventh and eight factors producing inconsistent loadings. Third, the visual examination of the scree plot additionally supported the retention of six factors. Fourth, there was a large drop off in observed score variance between the sixth and seventh factors. Finally, the first six factors showed clear variable groupings by activity area (e.g., Fitness activities, outdoor pursuits, etc.).

In the end, this PI chose to go against the recommendations of Horn's Parallel Analysis, and the eigenvalue > 1.0 rule (which supported 10 factors), to err on the side of logic, and the literature that supports the physical activity area groupings (AAHPERD 2103; Fairclough et al., 2002; NASPE 2007; 2008; 2009). The seventh and eighth factors were then removed, thus proceeding with the use of six-factor model.

Factor results. The final six-factor model analyzed with both Varimax and Oblim rotations. The most common rotation method in EFA research is the Varimax method (Field, 2009). Varimax is an orthogonal rotation method that produces independent factors that have no multicollinearity (high correlations) and minimizes the number of variables that have high loadings on each factor (Conway & Huffcutt, 2003). "Therefore it tries to load a smaller number or variables highly onto each factor resulting in more interpretable clusters of factors" (Field, 2009, p. 644). Due to the fact that both the Varimax and the Oblim rotations revealed the same basic simple structure, the Varimax model was used as the final model to help reduce multicollinearity. A similar factor structure was found in the research conducted by Fabrigar et al., (1999), who concluded that simple structure that was this evident in both oblique and oblim rotations can be considered “superior simple structure.”

The final six-factor model accounted for 68.7% of total observed score variance. The factor structure revealed all of the activity areas grouping individually, minus the Target activities and the Net/Wall activities, which factored together. All factor scores in the Varimax model were greater than .40, accept for one variable, *Frequency of participation in Net/Wall activities.* Due to this variable grouping with all other Net/Wall activities, and even with the lower factor score, the variable was retained to for purposes of content validity.

Five of the final six factors were named based on activity area grouping (i.e., all fitness activity items loaded together, thus naming the factor *self-efficacy to teach fitness activities*): (a) fitness activities, (b) outdoor pursuits, (c) individual performance activities, (d) dance/rhythmic activities, and (e) aquatic activities. Due to the fact that the *Net/Wall* and *Target* activity areas loaded together on the same factor (reducing seven activity areas, down to six), the new factor representing the grouping the two needed to be re-named. After further examination of the literature supporting lifetime physical activities in physical education (AAHPERD 2103; CDC 1997; Corbin, 2002; Fairclough, et al., 2002; NASPE 2007; 2008; 2009), this PI chose to rename the factor as *Hand/Eye Activities*, with potential future research addressing the reduction of these Hand/Eye variables.

The factor correlation matrix showed patterns that did have some relationship among variables. Tabachnick, et al., 2001 suggest that factor correlations >.32 should be considered for re-evaluation. Results showed moderate correlations between factors:

- 1 (Individual) and 4 (Aquatic) (.475)
- 1 (Individual) and 5 (Dance) (.356),
- 1 (Individual) and 6 (Fitness) (-.435)
- 3 (Outdoor) and 4 (Aquatic) (.328).

Correlations between the remaining factors were notably smaller (all < .27). These results show that factor correlations, especially with the Individual performance activity variables, share some small positive relationship among variables.

Reliability of Survey Instrument

RQ2: Does the Physical Educator Efficacy Scale for Teaching of Lifetime Physical Activities (PEES-LPA) provide evidence for reliability in measuring physical educator self-efficacy perceptions towards instruction of lifetime physical activities?

Internal Consistency. Reliability estimates for the full six-factor model revealed a high level of internal consistency, Cronbach's alpha = .95 (Nunnally, 1978). In addition, each of the six factors (activity groupings) were analyzed for item-total correlations and Cronbach's alpha if deleted. Results for all six factors were additionally quite high (alpha range .924-.955). Historically speaking, alpha scores greater than .70 are considered to be acceptable for scale development, with scores between .85-.95 deemed excellent (DeVellis, 2012; Hinkin, 1995, Nunnally, 1978).

Though all of the reliability scores, including the full model, all fall within the *excellent* category (Hinkin, 1995; Nunnally, 1978), these results should be interpreted cautiously. Clark and Watson (1995) address that there is such a thing as having internal consistency that is too high, thus having overly redundant items that are measuring content that is far too specific. They also address that high internal consistency can work against content validity, and that reliability scores only show that a portion of the construct is being measured, just over-and-over!

Additionally, some researchers (Clark and Watson, 1995; Field, 2009; Hinkin, 1995; Nunnally, 1978) have noted that internal consistency can be highly sensitive to the number of variables in an instrument. Therefore, it is possible to acquire a high internal consistency alpha because there are a lot of variables on an instrument, and not because all the items are highly reliable.

Inter-Item Correlations. Inter-item correlations were also evaluated to test the correlation between each item and the total score from the PEES-LPA (Hinkin, 1995). Inter-item correlations that are at the high extreme (> .85) suggested that a set of items are not contributing to something unique, and therefore are redundant. Additionally, variables that have a great deal of correlations below .30, then should be considered for removal.

Results from the inter-item correlations showed the following:

- The *Assess* and *Feedback* variables appeared to have a very high correlation
- The *Present* and *Identify* variables appeared to highly correlate

- The *Frequency of Participation* variables did not have strong correlations with many variables

The results from these correlations suggest that future research may look to combine variables, as well as removal of the *Frequency* variable.

Recommendations for Future Research

The results of this study provided support for the self-efficacy theory as the framework for measuring perceptions towards instruction of lifetime physical activities in physical education. It is suggested that future research should be conducted to address the following:

1. Test the full instrument using Confirmatory Factor Analysis (CFA) to assess unidimensionality of the construct.
2. Replicate the study with a larger, and more diverse sample that are not dependent on a third party (state AAHPERD) to help distribute the survey, thus trying to solicit the participants directly. The larger sample size will also grant future researchers the latitude to use-ad hoc deletion methods on univariate and multivariate outliers.
3. Use test-retest reliability procedures to test the current PEES-LPA
4. Conduct a focus group with the desired testing population (e.g., in-service physical educators)
5. If using an online survey, use a built in mechanism that will alert the participant if they skipped a question, thus attempting to reduce missing data
6. Validate the instrument, and factor structure, based on the reduction of factor variables

a. Reducing the items from Net/Wall and Target activities into a new activity area: Hand/Eye activities

b. Possible removal (or re-evaluation) of the Individual Performance activity variables due to the moderate factor correlation it had with three other factor groupings.

7. Validate the item by reducing items that have inter-item correlation extremes:

   a. Though much of Bandura's (1982; 1997) research emphasizes the importance mastery experiences have on self-efficacy, the PEES-LPA items asking: "In the last year, how often did you participate…" appeared not to have a strong influence on those self-efficacy perceptions. This PI recommends deletion of those items.

   b. The *Feedback* and *Assessment* items had very high extremes, and thus a recommendation is made to combine these two items. The combination of these variables is literature supported, and would be a logical area for item reduction, especially due to the fact that skill-specific feedback is a form of assessment (AAHPERD 2103; Fairclough, et al., 2002; NASPE 2007; 2008; 2009).

   c. The *Present* and *Identify* (of critical skill elements) items also showed high correlation. Further literature examination should take place to evaluate the relationship between these two variables, leading to possible item reduction

8. Investigate the accuracy of self-reported teacher behavior in regard to instruction of lifetime physical activities (as well as other sport-specific critical skill elements.

A hypothesized re-designed PEES-LPA instrument would then have five activity areas (Fitness, Outdoor Pursuits, Aquatics, Dance/Rhythmic, and Individual Sports), opposed to the seven used in this study. Individual sports would represent the absorption of: (a) *Net/Wall* activities, (b) *Target* activities, and (c) *Individual Performance* activities. The revised instrument would include six self-efficacy questions, opposed to the nine used in this study (removal of *Frequency of Participation*, confidence to *Assess* critical skill elements, and confidence to *Identify and Define* critical skill elements). Overall, this hypothesized instrument would then have 30 total items, opposed to the 63 used in this study.

It is important to note here that these recommendations should not be taken strictly at face value, primarily due to the sample of participants being small in size and not representative of the full range of physical educators teaching in the US. For example, it was clear that due to: (a) the high average of years teaching physical education (M = 15.2), (b) the high self-efficacy scores (especially in fitness, net/wall, and target activities) (see Appendix K), and (c) research bias towards professionally acute physical educators, the sample was less than ideal for representing physical educators across the US. Thus, re-examining of the content and construct validity through further literature examination and expert review is recommended prior to survey item reduction.

Conclusion

In conclusion, the growing body of evidence and support for the promotion of lifetime physical activities through school-based physical education is more vital than ever. Numerous government and national initiatives have strongly emphasized that lifetime physical activities should be at the forefront of secondary physical education programs in the US. As a result of the value placed on these physical education programs, it becomes important to examine how confident physical educators are to deliver such skills and competencies.

As a result of no current instrument measuring self-efficacy perceptions of physical educators towards teaching lifetime physical activities, this study employed systematic procedures to begin developing a scale to fill the aforementioned void. The current study offers preliminary support for the psychometric properties of the PEES-LPA for validity and reliability. The major findings from this study demonstrated that there was strong evidence for both content and construct validity, this was evident due to: (a) resulting factors showing simple structure that aligns with literature supporting the classification of lifetime physical activities (AAHPERD, 2013), (b) factors were composed of items that logically relate, (c) internal consistency showed to be very high, though should be interpreted with caution due to the high number of survey items as well as the presence of high inter-item correlations, and (d) the PEES-LPA appears to be an appropriate instrument for measuring self-efficacy perceptions of physical educators, though further revisions would be needed to help remove possible multicollinearity issues.

The PEES-LPA is the first instrument to measure self-efficacy perceptions of teachers towards the instruction of lifetime physical activities. The results of this study indicated that self-efficacy theory was an appropriate means for measuring perceptions towards instruction, using diverse teaching task demands. Further study should be conducted to reduce the number of variables to help reduce correlation extremes, and confirm the findings of the current study.

References

Allinder, R. M. (1994). The relationship between efficacy and the instructional practices of special education teachers and consultants. *Teacher Education and Special Education*, *17*, 86-95.

American Alliance for Health, Physical Education, Recreation, & Dance (AAHPERD). (2013). *National standards & grade level outcomes in K-12 physical education*. Reston, VA; author.

American College of Sports Medicine. (2011). *ACSM's complete guide to fitness & health.* Champaign, IL: Human Kinetics.

American Educational Research Association (AERA). (1999). *Standards for Educational and Psychological Testing*. Washington, D.C.: author.

Amusa, L. O., Toriola, A. L., & Goon, D. T. (2012). Youth, physical activity and leisure education: Need for a paradigm shift. *African Journal for Physical and Health Education, Recreation, and Dance*, *18* (4), 992–1006.

Ashton, P., & Webb, R. B. (1986). *Making a difference: Teachers' sense of efficacy and student achievement*. New York: Longman.

Australian Institute of Health and Welfare (2004). *Australia's health*. Bouchard, Canberra: Australian Institute of Health and Welfare.

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavior change. *Psychological Review, 84*, 191–215.

Bandura, A. (1982). Self-efficacy mechanism in human agency. *American Psychologist, 37*, 122-147.

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

Bandura, A. (1993). Perceived self-efficacy in cognitive development and functioning. *Educational Psychologist, 28*, 117-148.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: Freeman.

Bandura, A. (2001). Social cognitive theory: An agentive perspective. *Annual Review of Psychology*, *52*, 1–26.

Bandura, A. (2006). Guide for constructing self-efficacy scales. In F. Pajares & T. Urdan (Eds.), *Self-efficacy beliefs of adolescents* (Vol. 5, pp. 307–337). Greenwich, CT: Information Age Publishing.

Barrett, G. V. (1972). Models of the future for industrial and organizational psychology. *Personnel Psychology*, *25*(1), 1-17.

Bell, J., Rogers, V. W., Dietz, W. H., Ogden, C. L., Schuler, C., & Popovic, T. (2011). CDC grand rounds: Childhood obesity in the United States. *MMWR Morb Mortal Wkly Rep*, *60*, 42-46.

Biddle, S., & Goudas, M. (1998). Physical education teacher efficacy: Scale development and relationship with curricular goals. *Journal of Sports Science*, *19*, 23-32.

Block, M. E., Hutzler, Y., Barak, S., & Klavina, A. (2013). Creation and validation of the self-efficacy instrument for physical education teacher education majors toward inclusion. *Adapted Physical Activity Quarterly*, *30*(2), 184-205.

Boyce, B. A., & Rikard, G. L. (2011). A Comparison of Supply and Demand for PETE Professionals in Higher Education in the United States. *Journal of Teaching in Physical Education*, *30(2),* 103-115.

Burton, L. J., & Mazerolle, S. M. (2011). Survey instrument validity part I: Principles of survey instrument development and validation in athletic training *Education Research*, *6* (1), 27–35.

Carroll, B. & Loumidis, J. (2001). Children's perceived competence and enjoyment in physical education and physical activity outside school. *European Physical Education Review, 7* (1), 24-43.

Caspersen C. J., Powell K. E., & Christenson G. M. (1985). Physical activity, exercise, and physical fitness: Definitions and distinctions for health-related research. *Public Health, 100* (2), 126-31.

Caprara G. V., Barbaranelli C., & Borgogni L., Pettita, L., & Rubinacci, A. (2003). Teachers', school staff's and parents' efficacy beliefs as determinants of attitudes toward school. *European Journal of Psychology of Education, 18* (1), 15-31.

Centers for Disease Control and Prevention-CDC. (1997). Guidelines for school and community programs to promote lifelong physical activity among young people. *Morbidity & Mortality Weekly Report, 46*, 1–36.

Centers for Disease Control and Prevention. (2010). *State indicator report on physical activity, 2010*. Atlanta, GA: U.S. Department of Health and Human Services.

Centers for Disease Control and Prevention. (2012). Youth risk behavior surveillance-United States, 2011. *Morbidity & Mortality Weekly Report, 61*, 1-162.

Child, D. (2006). *The essentials of factor analysis*. (3rd ed.). New York: Bloomsbury.

Clark, L. A., & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment*, *7* (3), 309.

Comrey, A. L. (1988). Factor-analytic methods of scale development in personality and clinical psychology. *Journal of consulting and clinical psychology*, *56*(5), 754.

Conway, J. M., & Huffcutt, A. I. (2003). A review and evaluation of exploratory factor analysis practices in organizational research. *Organizational Research Methods*, *6*(2), 147-168.

Corbin, C. B. (2002). Physical activity for everyone: What every physical educator should know about promoting lifelong physical activity. *Journal of Teaching in Physical Education*, *21*(2), 128-44.

Costello, A. B., & Osborne, J. W. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research, & Evaluation, 10*, 1-9

Creswell, J. W. (2013). *Research design: Qualitative, quantitative, and mixed methods approaches*. London: Sage Publications.

Cronbach, L. (1951). Coefficient alpha and the internal structures of tests. *Psychometrika*, 16, 297-334.

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

Cronbach, L. J. (1971). Test validation. *Educational Measurement*, *2*, 443-507.

DeCoster, J. (2005). *Scale construction notes*. Retrieved September 7, 2013 from http://www.stat-help.com/notes.html

Dellinger, A. B., Bobbett, J. J., Olivier, D. F., & Ellett, C.D. (2008). Measuring teachers' self-efficacy beliefs: Development and use of the TEBS-Self. *Teaching and Teacher Education: An International Journal of Research Studies, 24*(3), 751-766.

DeVellis, R. F. (2012). *Scale development: Theory and applications (3rd).* London: Sage Publications

Downing, S. M. (2003). Validity: On the meaningful interpretation of assessment data. *Medical Education*, *37*(9), 830-837.

Edginton, C. R., Kirkpatrick, B., Schupbach, R., Phillips, C., & Chen, M. C. P. (2011). A dynamic pedagogy of physical education teacher preparation□: Linking practice with theory. *Asian Journal of Physical Education and Recreation 16*(2), 7–25.

Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, *4*(3), 272.

Fairclough, S., Stratton, G., & Baldwin, G. (2002). The contribution of secondary school physical education to lifetime physical activity. *European Physical Education Review*, *8*(1), 69–84.

Felt, D. L., Short, S. E., & Sullivan, P. J. (2008). *Self-efficacy in sport: Research and strategies for working with athletes, teams and coaches*. Champaign, IL: Human Kinetics.

Field, A. (2009). *Discovering statistics using SPSS*. London: Sage publications.

Ford, J. K., MacCallum, R. C., & Tait, M. (1986). The application of exploratory factor analysis in applied psychology: A critical review and analysis. *Personnel Psychology*, *39*(2), 291-314.

Fowler, F. J. (2009). *Survey research methods*. London: Sage Publications.

Fulcher, G. (1999). Assessment in English for academic purposes: Putting content validity in its place. *Applied Linguistics, 20*(2), 221-236.

Gencay, O. A. (2009). Validation of the physical education teachers' physical activity self-efficacy scale with a Turkish sample. *Social Behavior and Personality: An International Journal*, *37*(2), 223–230.

Ghiselli, E. E., Campbell, J. P., & Zedeck, S. (1981). *Measurement theory for the behavioral sciences* (Vol. 8). San Francisco: WH Freeman.

Gibson, S., & Dembo, M. (1984). Teacher efficacy: A construct validation. *Journal of Educational Psychology, 76*, 569-582

Gist, M. E., & Mitchell, T. R. (1992). Self-efficacy: A theoretical analysis of its determinants and malleability. *Academy of Management review*, *17*(2), 183-211.

Gordon-Larsen, P., McMurray, R. G., & Popkin, B. M. (2002). Determinants of adolescent physical activity and inactivity patterns. *Pediatrics*, *105*(6), 83-83.

Gorozidis, G., & Papaioannou, A. (2011). Teachers' self-efficacy, achievement goals, attitudes and intentions to implement the new Greek physical education curriculum. *European physical education review*, *17*(2), 231-253.

Gorsuch, R. L. (1983). Factor analysis (2nd ed.). Hillsdale, NJ:Erlbaum.

Groves, R. M., Fowler Jr, F. J., Couper, M. P., Lepkowski, J. M., Singer, E., & Tourangeau, R. (2013). *Survey methodology*. Hoboken, New Jersey: Wiley.

Guadagnoli, E., & Velicer, W. F. (1988). Relation to sample size to the stability of component patterns. *Psychological Bulletin*, *103*(2), 265.

Guskey, T. R. (1988). Teacher efficacy, self-concept, and attitudes toward the implementation of instructional innovation. *Teaching and Teacher Education,* *4*, 63-69.

Haeusner, M. K. (1997). *A survey investigating the self-efficacy levels of Wisconsin adapted physical education teachers*. (Doctoral dissertation). Retrieved form ProQuest on February 21, 2014 from http://search.proquest.com.proxy.its.virginia.edu/dissertations/docview/304395685/previewPDF/75EE4A74CD64C09PQ/1?accountid=14678

Hall, B., Burley, W., Villeme, M., & Brockmeier, L. (1992). *An attempt to explicate teacher efficacy beliefs among first year teachers.* Paper presented at the annual meeting of the American Educational Research Association, San Francisco.

Haynes, S. N., Richard, D., & Kubany, E. S. (1995). Content validity in psychological assessment: A functional approach to concepts and methods. *Psychological Assessment*, 7(3), 238.

Henson, R. K., Kogan, L. R., Vacha-Haase, T. (2001). A reliability generalizations study of the teacher efficacy scale and related instruments. *Educational and Psychological Measurement, 36* (3). 404-420.

Hillsdon, M., & Foster, C. (2003). The relationship between physical activity during childhood and adolescence and coronary heart disease risk factors in young adulthood. In A. Giles (Ed.), *A Lifecourse Approach to Coronary Heart Disease Prevention: Scientific and Policy Revi*ew, (pp. 199–212). London: The Stationery Office.

Hinkin, T. (1995). A review of scale development practices in the study of organizations. *Journal of Management*, *21*(5), 967–988.

Hinkin, T. (1998). A brief tutorial on the development of measures for use in survey questionnaires. *Organizational Research Methods*, *1*(1), 104-121.

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika, 30*, 179-185.

Hox, J. J. (2012). From theoretical concept to survey question. In Lyberg, L. E., Biemer, P., Collins, M., De Leeuw, E. D., Dippo, C., Schwarz, N., & Trewin, D. (Eds.), *Survey Measurement and Process Quality*. New York, NY; Wiley-Interscience.

Hoy, W. K., & Woolfolk, A. E. (1993). Teachers' sense of efficacy and the organizational health of schools. *The Elementary School Journal, 93*, 356-372.

Huitt, W., & Hummel, D. K. (1999). *Internal and External Validity: General Issues*. Retrieved March 4, 2014 from http://www.edpsycinteractive.org/topics/intro/valdgn.html.

Humphries, C. A., Hebert, E., Daigle, K., & Martin, J. (2012). Development of a physical education teaching efficacy scale. *Measurement in Physical Education and Exercise Science*, *16*(4), 284–299.

Hurley, A. E., Scandura, T. A., Schriesheim, C. A., Brannick, M. T., Seers, A., Vandenberg, R. J. (1997). Exploratory and confirmatory factor analysis: Guidelines, issues, and alternatives. *Journal of Organizational Behavior, 18*, 667-683.

Janssen, I., & LeBlanc, A. G. (2010). Review systematic review of the health benefits of physical activity and fitness in school-aged children and youth. *International Journal of Behavioral Nutrition and Physical Activity*, *7*(40), 1-16.

Johnson, J. (2011, May). *What do we measure? Methodological versus institutional validity in student surveys*. St. Louis: Association for Institutional Research Annual Forum.

Kaiser, H. F. (1958). The application of electronic computers to factor analysis. *Educational and Psychological Measurement, 20*, 141–151.

Krosnick, J. A., & Fabrigar, L. R. (1997). Designing rating scales for effective measurement in surveys. *Survey Measurement and Process Quality*, *50*, 141-164.

Kujala, U. M., Kaprio, J., Sarna, S., & Koskenvuo, M. (1998). Relationship of leisure-time physical activity and mortality. *JAMA: The Journal of the American Medical Association*, *279*(6), 440-444.

Lamarche, L., Gammage, K. L., Sullivan, P. J., & Gabriel, D. A. (2013). A Psychometric Evaluation of the Self-Presentational Efficacy Scale. *Measurement in Physical Education and Exercise Science*, *17*(2), 120–134.

Lavrakas, P. J. (2008). *Encyclopedia of survey research methods*. Washington DC: SAGE Publications, Incorporated.

Leedy, P. D., & Ormrod, J. E. (2001). Practical research planning and design, 7th Ed., Prentice-Hall, Upper Saddle River, N.J.

Le Masurier, G., & Corbin, C. B. (2006). Top 10 reasons for quality physical education. *Journal of Physical Education, Recreation & Dance*, *77*(6), 44-53.

Little, R. J. (1988). A test of missing completely at random for multivariate data with missing values. *Journal of the American Statistical Association*, *83*(404), 1198-1202.

Litwin, M. S. (1995). *How to measure survey reliability and validity* (Vol. 7). New York: Sage.

Lubans, D., & Morgan, P. (2008). Evaluation of an extra-curricular school sport program promoting lifestyle and lifetime activity for adolescents. *Journal of Sports Sciences*, *26*(5), 519–29.

Lucko, G., & Rojas, E. M. (2009). Research validation: Challenges and opportunities in the construction domain. *Journal of Construction Engineering and Management, 136*(1), 127-135.

Mackenzie, S. B., Podsakoff, P. M., & Podsakoff, N. P. (2011). Construct measurement and validation procedures in MIS and behavioral research: Integrating new and existing techniques. *MIS Quarterly*, *35*(2), 293–334.

Maddux, J. E. (1995). *Self-efficacy, adaptations and adjustment: Theory, research, and application*. New York: Plenum Press.

Malinen, O. P., Savolainen, H., & Xu, J. (2013). Dimensions of teacher self-efficacy for inclusive practices among mainland Chinese pre-service teachers. *Journal of International Special Needs Education*, *16*(2), 82-93.

Martin, J. J., & Kulinna, P. H. (2003). The development of a physical education teachers ' physical activity self-efficacy instrument. *Journal of Teaching Physical Education, 22,* 219–232.

Martin, J. J., & Kulinna, P. H. (2005). A social cognitive perspective of physical-activity-related behavior in physical education. *Journal of Teaching in Physical Education, 24*, 265–281.

McKinnon, J. (1988). Reliability and validity in field research: Some strategies and tactics. *Accounting, Auditing & Accountability Journal*, *1*(1), 34-54.

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.

Nader, P. R., Bradley, R. H., Houts, R. M., McRitchie, S. L., & O'Brien, M. (2008). Moderate-to-vigorous physical activity from ages 9 to 15 years. *The Journal of the American Medical Association*, *300*(3), 295-305.

National Association for Sport and Physical Education (NASPE). (2008). *National standards & guidelines for physical education teacher education, (3rd ed.*). Reston, VA: author.

National Association for Sport and Physical Education. (2007). *Physical education teacher evaluation tool, (3rd ed.).* Reston, VA: author.

National Association for Sport and Physical Education. (2009). *Moving into the future: National standards for physical education, (3rd ed.).* Reston, VA: author.

National Council for Accreditation of Teacher Education (NCATE). (2008). *Professional standards for the accreditation of schools, colleges, and departments of education*. Washington, DC: author.

Nevo, B. (1985). Face validity revisited. *Journal of Educational Measurement*, *22*(4), 287-293.

Nunnally, J. (1978). *Psychometric theory*. McGraw-Hill: New York

Ntoumanis, N., & Biddle, S. J.H. (1999). A review of motivational climate in physical activity. *Journal of Sports Sciences, 17*, 643-665.

Ogden, C. L., Carroll, M. D., Kit, B. K., & Flegal, K. M. (2012). Prevalence of obesity and trends in body mass index among US children and adolescents, 1999-2010. *Journal of the American Medical Association. 307*(5):483-490.

Pan, Y. (2012). The development of a teachers' self-efficacy instrument for high school physical education teacher. *World Academy of Science, Engineering, and Technology, 66*, 1152–1157.

Pajares, F. (1996). Self-efficacy beliefs in academic settings. *Review of Educational Research, 66*, 533-578.

Pajares. F. (2002). *Overview of social cognitive theory and of self-efficacy*. Retrieved May 05, 2013 from http://www.uky.edu/~eushe2/Pajares/eff.html.

Peng, C. Y. J., Harwell, M., Liou, S. M., & Ehman, L. H. (2006). Advances in missing data methods and implications for educational research. *Real Data Analysis*, 31-78.

Pollock, M. L., & Wilmore, J. H. (1990). *Exercise in health and disease: Evaluation and prescription for prevention and rehabilitation, (2nd ed.),* Philadelphia: Saunders W. B.

Pushka, P., Benaziza, H., & Porter, D. (2003). *Physical activity and health*. Geneva, Switzerland: World Health Organization.

Ross, J. A. (1992). Teacher efficacy and the effect of coaching on student achievement. *Canadian Journal of Education, 17*(1), 51-65.

Ross, J. A. (1994, June). *Beliefs that make a difference: The origins and impacts of teacher efficacy*. Paper presented at the annual meeting of the Canadian Association for Curriculum Studies, Alberta.

Ross, J. G., Dotson, C.O., Gilbert, G. G. and Katz, S. J. (1985). What are kids doing in school physical education? *Journal of Physical Education, Recreation and Dance, 56*(1): 73–76.

Ross, S., Francis, L. A., BeLue, R. Z., & Viruell-Fuentes, E. A. (2012). Associations between physical activity and overweight among U.S. youth by immigrant generation: Results from the 2007 National Survey of Children's Health. *Journal of Physical Activity & Health*, *9*(6), 840-848.

Russell, D. W. (2002). In search of underlying dimensions: The use (and abuse) of factor analysis. *Personality and Social Psychology Bulletin*, *28*(12), 1629-1646.

Sallis, J. F., Simons-Morton, B. G., Stone, E. J., & Corbin, C. B. (1992). Determinants of physical activity and interventions in youth. *Medicine & Science in Sports & Exercise*. *24*8-257.

Scrabis-Fletcher, K., & Silverman, S. (2010). Perception of competence in middle school physical education: Instrument development and validation. *Research Quarterly for Exercise and Sport*, *81*(1), 52-61.

Schlomer, G. L., Bauman, S., & Card, N. A. (2010). Best practices for missing data management in counseling psychology. *Journal of Counseling Psychology*, *57*(1), 1.

Schwartz, R. M., & Gottman, J. M. (1976). Toward a task analysis of assertive behavior. *Journal of Consulting and Clinical Psychology*, *44*(6), 910.

Snyder, T. D., Dillow, S. A. (2012) *Digest of Education Statistics 2011.* National Center for Education Statistics, Institute of Education Sciences, Washington, DC: US Department of Education

Stajkovic, A. D., & Luthans, F. (1998). Self-efficacy and work-related performance: A meta-analysis. *Psychological Bulletin*, *124*(2), 240.

Stevens, J. P. (2009). *Applied multivariate statistics for the social sciences*. New York, N.Y.: Taylor & Francis.

Stein, M. K., & Wang, M. C. (1988). Teacher development and school improvement: The process of teacher change. *Teaching and Teacher Education, 4*, 171-187.

Streiner, D. L., & Norman G. R. (1989). *Health measurement scales: A guide to their development and use*. Oxford: Oxford University Press.

Strong, W. B., Malina, R. M., Blimkie, C. J., Daniels, S. R., Dishman, R. K., Gutin, B.

(2005). Evidence based physical activity for school-age youth. *Journal of Pediatrics, 146*, 732 – 737.

Swisher, L. L., Beckstead, J. W., & Bebeau, M. J. (2004). Factor analysis as a tool for survey analysis using a professional role orientation inventory as an example. *Physical Therapy*, *84*(9), 784-799.

Tabachnick, B. G., Fidell, L. S., & Osterlind, S. J. (2001). Using multivariate statistics. New York: Pearsons.

Taliaferro, A. (2011). *Validating an instrument to investigate physical educators' self-efficacy toward inclusion*. (Doctoral dissertation). Retrieved on November 21, 2013 from http://search.proquest.com.proxy.its.virginia.edu/dissertations/docview/822408625/E923F06A70854965PQ/1?accountid=14678.

Taylor, S., & Asmundson, G. J. (2008). Internal and external validity in clinical research. *Handbook of Research Methods in Abnormal and Clinical Psychology*, *1*, 23-34.

Tinsley, H. E., & Tinsley, D. J. (1987). Uses of factor analysis in counseling psychology research. *Journal of Counseling Psychology*, *34*(4), 414.

Thorndike, R. M. (1997). *Measurement and evaluation in psychology and education* (6th ed.). New Jersey: Prentice Hall, Inc.

Thorndike, R. M. (2005). *Measurement and evaluation in psychology and education* (7th ed.). New Jersey: Pearson Education, Inc. Tinsley.

Treasure, D. C., & Roberts, G. C. (2001). Students' perceptions of the motivational climate, achievement beliefs, and satisfaction in physical education. *Research*

Quarterly for Exercise and Sport, 72, 165-175.

Troiano, R. P., Berrigan, D., Dodd, K. W., Mâsse, L. C., Tilert, T., & McDowell, M. (2008). Physical activity in the United States measured by accelerometer. *Medicine & Science in Sport & Exercise, 40* (1), 181-188.

Tschannen-Moran, M., Woolfolk Hoy, A. & Hoy, W. K. (1998). Teacher efficacy: Its meaning and measure. *Review of Educational Research, 68*, 202 - 248.

Tschannen-Moran, M., & Woolfolk-Hoy, A. (2001). Teacher efficacy: capturing an elusive construct. *Teaching and Teacher Education, 17,* 783-805.

Umhoeffer, D. L. (2011). *Adapted physical education service models and the effects on the perceived efficacy of general physical education teachers*. (Doctoral dissertation). Retrieved on November 21, 2013 from http://search.proquest.com.proxy.its.virginia.edu/dissertations/docview/868328346/62298C3CE4B84A62PQ/1?accountid=14678.

U. S. Depart of Health and Human Services (USDHHS). (2008). *Physical activity guidelines for Americans: Be active, healthy and happy!.* Washington DC: author.

U. S. Depart of Health and Human Services- USDHHS, Center for Disease Control, National Center for Health Statistics. (2011). *Health, United States, 2011: With special feature on socioeconomic status and health*. Washington DC: author.

U.S. Department of Health and Human Services- USDHHS. (2012). *Physical Activity Guidelines for Americans Midcourse Report: Strategies to Increase Physical Activity Among Youth*. Washington, DC: author.

Vacha-Haase, T. (1998). Reliability generalization: Exploring variance in measurement error affecting score reliability across studies. *Educational and Psychological Measurement*, *58*, 6-20.

Velicer, W. F., & Fava, J. L. (1998). Affects of variable and subject sampling on factor pattern recovery. *Psychological Methods*, *3*(2), 231.

Warburton, D. E., Nicol, C. W., & Bredin, S. S. (2006). Health benefits of physical activity: The evidence. *Canadian Medical Association Journal*, *174*(6), 801-809.

Wallhead, T. L., & Buckworth, J. (2004). The role of physical education in the promotion of youth physical activity. *Quest*, *56*(3), 285–301.

Ward, R. A. (2005). Impact of mentoring on teacher efficacy. *Academic Exchange Quarterly, Winter*, 148-154.

Appendix A

IRB Approval

UNIVERSITY of VIRGINIA

OFFICE OF THE VICE PRESIDENT FOR RESEARCH
INSTITUTIONAL REVIEW BOARD FOR THE SOCIAL AND BEHAVIORAL SCIENCES

In reply, please refer to: Project # 2014-0020-00

February 12, 2014

Kason O'Neil
Barbara Boyce
Human Services
7297 Jackson Ave., Apt. G
Mechanicsville, VA 23111

Dear Kason O'Neil and Barbara Boyce:

The Institutional Review Board for the Behavioral Sciences has approved your research project entitled "Survey on the Status of Lifetime Physical Activities in Secondary Physical Education." You may proceed with this study.

This project # 2014-0020-00 has been approved for the period February 12, 2014 to February 11, 2015. If the study continues beyond the approval period, you will need to submit a continuation request to the Review Board. If you make changes in the study, you will need to notify the Board of the changes.

Sincerely,

Tonya R. Moon, Ph.D.
Chair, Institutional Review Board for the Social and Behavioral Sciences

One Morton Drive, Suite 500 • Charlottesville, VA 22903
P.O. Box 800392 • Charlottesville, VA 22908-0392
Phone: 434-924-5999 • Fax: 434-924-1992
www.virginia.edu/vpr/irb/sbs.html

Appendix B

Letter to Request Expert Review

Dear Dr. XXXXX,

My name is Kason O'Neil and I am a doctoral student at the University of Virginia studying Kinesiology (Physical Education Pedagogy) under the advisement of Dr. Ann Boyce.

Due to your expertise in PE pedagogy, and your previous cooperation with the 2012 research I helped conduct with Dr. Boyce & Dr. Jackie Lund, I am seeking your expert feedback on study I am conducting examining the self-efficacy perceptions of secondary physical educators towards the instruction of lifetime physical activities. My study consists of the development and validation of a new self-efficacy instrument.

If you are willing to participate with the expert review process of this newly constructed instrument, please read over the directions in the attached word document, which will then prompt you to fill out the expert review form. This expert review form should take 10-15 minutes of your time. Upon completion, save your answers as a word or pdf document and send the completed form back to me at kmo7rh@virginia.edu.

If you have any questions, please feel free to contact me via email or phone at XXX-XXX-XXXX If you have a question specifically for Dr. Boyce, she can be contacted at bab6n@virginia.edu.

Thank your for your time and consideration,

Kason O'Neil, M.S.S.
Doctoral Candiate- Kinesiology
University of Virginia

Appendix C

Expert Review Evaluation Sheet

Dear Professor,

First off, I would like to thank you for your time and willingness to participate in the expert review process for the creation and validation of the *Physical Educator Efficacy Scale for Teaching Lifetime Physical Activities (PEES-LPA).* During the early stages of development for this scale, your feedback is highly valued.

In 2013, AAHPERD published the *National Standards & Grade Level Outcomes in K-12 Physical Education*. This publication operationally defined lifetime physical activity and delimited seven specific activity categories (with associated activities) that meet the defining characteristics for being a lifetime physical activity. The seven lifetime physical activity (LPA) categories specified by AAHPERD (2013) are:

1. Outdoor Pursuits (e.g., hiking, orienteering, rock climbing)
2. Fitness Activities (e.g., strength/weight training, yoga, cycling/biking)
3. Aquatic Activities (e.g., swimming, diving)
4. Individual Sports (e.g., track & field, self defense, gymnastics)
5. Dance/Rhythmic Activities (e.g., creative, line, modern)
6. Net/Wall Activities (e.g., tennis, badminton, pickleball)
7. Target Activities (e.g., golf, bowling, archery)

The purpose of this study is to create a scale that examines how confident (self-efficacious) physical educators are to deliver the skills and competencies necessary to engage students in each of the seven LPA categories. The task constraints used in the PEE-LPA, that would most influence self-efficacy perceptions towards instruction of LPA, were selected from two NASPE resources: *National Standards & Guidelines for Physical Education Teacher Education* (2008), & the *Physical Education Teacher Evaluation Tool* (2007).

Directions: Your participation in this expert review will involve observing PEES-LPA items (extracted from an online survey website) and provide feedback addressing participant (1) demographics, (2) personal experience, and (3) self-efficacy towards instruction. Please fill out the following form at your earliest convince. Upon completion, save the file (as a word or PDF document) and send the completed form to **Kason O'Neil at kmo7rh@virginia.edu.**

Thank you again for your time and consideration,

Kason O'Neil, M.S.S
Doctoral Candidate- Kinesiology (Physical Education Pedagogy)
University of Virginia

LIFETIME PHYSICAL ACTIVITY CATEGORIES

Outdoor Pursuits (e.g., hiking, orienteering, rock climbing)
Fitness Activities (e.g., strength/weight training, yoga, cycling/biking)
Aquatic Activities (e.g., swimming, diving)
Individual Sports (e.g., track & field, self defense, gymnastics)
Dance/Rhythmic Activities (e.g., creative, line, modern)
Net/Wall Activities (e.g., tennis, badminton, pickleball)
Target Activities (e.g., golf, bowling, archery)

1. The activity categories above are clear and concise.
   ☐ Strongly Disagree ☐ Disagree ☐ Neutral ☐ Agree ☐ Strongly Agree

2. In regards to physical education, the activity categories above are accurate in representing the operational definition of *Lifetime Physical Activities.*
   ☐ Strongly Disagree ☐ Disagree ☐ Neutral ☐ Agree ☐ Strongly Agree

3. (Optional) Please provide any overall feedback you may have on the lifetime physical activity categories:

PEES-LPA SECTION 1- DEMOGRAPHIC INFORMATION

1. Which grade(s) do you currently teach secondary physical education?

- ○ Middle School (5-8)
- ○ High School (9-12)
- ○ Both HS & MS
- ○ Non of the above

2. Which state are you currently certified to teaching physical education?

-- Select --

3. How many years of experience do you have as a certified physical education teacher?

4. What is your highest education degree?

- ○ Bachelors
- ○ Masters
- ○ Doctorate

5. Gender

- ○ Male
- ○ Female

6. Please provide your email address in the space below.

Please answer the following questions based on the above PEES-LPA demographic questions:

4. The demographic questions above are clear and concise.

   ☐ Strongly Disagree ☐ Disagree ☐ Neutral ☐ Agree ☐ Strongly Agree

5. The format is appropriate for the selected demographic questions.

   ☐ Strongly Disagree ☐ Disagree ☐ Neutral ☐ Agree ☐ Strongly Agree

6. (Optional) Please provide any overall feedback you may have on:
   a. Question order
   b. Content of questions
   c. Question relevancy

Expert Response:

PEES-LPA SECTION 2- PERCEPTION OF EXPERIENCE WITH LIFETIME PHYSICAL ACTIVITIES

<u>Note to experts</u>: Section 2 of the PEES-LPA looks to examine perceptions educators have based on their experience with the different lifetime physical activity categories.

7. Please rate how you would classify your personal skill-ability within each activity category.

	Novice	Advanced Beginner	Competent	Proficient	Expert
Outdoor Pursuits (e.g., hiking, orienteering, rock climbing)	○	○	○	○	○
Fitness Activities (e.g., strength/weight training, yoga, cycling/biking)	○	○	○	○	○
Aquatic Activities (e.g., swimming, diving)	○	○	○	○	○
Individual Sports (e.g., track & field, self defense, gymnastics)	○	○	○	○	○
Dance/Rhythmic Activities (e.g., creative, line, modern)	○	○	○	○	○
Net/Wall Activities (e.g., tennis, badminton, pickleball)	○	○	○	○	○
Target Activities (e.g., golf, bowling, archery)	○	○	○	○	○

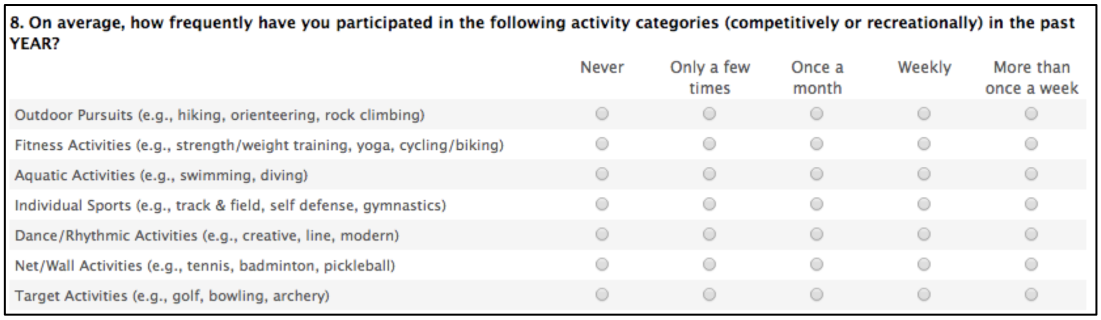
7. PEES-LPA question 7 is clear and concise.

   ☐ Strongly Disagree ☐ Disagree ☐ Neutral ☐ Agree ☐ Strongly Agree

8. The qualitative descriptors selected (e.g., novice, expert) are appropriate for measuring personal skill ability.

   ☐ Strongly Disagree ☐ Disagree ☐ Neutral ☐ Agree ☐ Strongly Agree

9. (Optional) Please provide any overall feedback you may have about question 7.



8. On average, how frequently have you participated in the following activity categories (competitively or recreationally) in the past YEAR?

	Never	Only a few times	Once a month	Weekly	More than once a week
Outdoor Pursuits (e.g., hiking, orienteering, rock climbing)	○	○	○	○	○
Fitness Activities (e.g., strength/weight training, yoga, cycling/biking)	○	○	○	○	○
Aquatic Activities (e.g., swimming, diving)	○	○	○	○	○
Individual Sports (e.g., track & field, self defense, gymnastics)	○	○	○	○	○
Dance/Rhythmic Activities (e.g., creative, line, modern)	○	○	○	○	○
Net/Wall Activities (e.g., tennis, badminton, pickleball)	○	○	○	○	○
Target Activities (e.g., golf, bowling, archery)	○	○	○	○	○

10. PEES-LPA question 8 is clear and concise.

☐ Strongly Disagree ☐ Disagree ☐ Neutral ☐ Agree ☐ Strongly Agree

11. The qualitative descriptors selected (e.g., once a month, weekly) are appropriate for measuring frequency of personal participation.

☐ Strongly Disagree ☐ Disagree ☐ Neutral ☐ Agree ☐ Strongly Agree

12. (Optional) Please provide any overall feedback you may have about question 8.

9. Compared to the average certified physical educator, please rate how you would classify your competence to teach a variety of activities within each activity category.

	Novice	Advanced Beginner	Competent	Proficient	Expert
Outdoor Pursuits (e.g., hiking, orienteering, rock climbing)					
Fitness Activities (e.g., strength/weight training, yoga, cycling/biking)					
Aquatic Activities (e.g., swimming, diving)					
Individual Sports (e.g., track & field, self defense, gymnastics)					
Dance/Rhythmic Activities (e.g., creative, line, modern)					
Net/Wall Activities (e.g., tennis, badminton, pickleball)					
Target Activities (e.g., golf, bowling, archery)					

13. PEES-LPA question 9 is clear and concise.

☐ Strongly Disagree ☐ Disagree ☐ Neutral ☐ Agree ☐ Strongly Agree

14. The qualitative descriptors selected (e.g., novice, expert) are appropriate for measuring perception of teaching competence.

☐ Strongly Disagree ☐ Disagree ☐ Neutral ☐ Agree ☐ Strongly Agree

15. (Optional) Please provide any overall feedback you may have about question 9.

PEES-LPA SECTION 3- SELF-EFFICACY TOWARDS INSTRUCTION OF LPA

<u>Note to experts</u>: Section 3 of the PEES-LPA looks to examine self-efficacy perceptions educators have towards task constraints associated with instruction of lifetime physical activities. These questions assess confidence levels based on Bandura's 0-10 model for measuring self-efficacy perceptions.

10. How confident are you in your ability to <u>IDENTIFY AND DEFINE</u> critical skill elements in the following activity categories?

	0– No Confidence	1	2	3	4	5– Moderately Confident	6	7	8	9	10 – Very Confident
Outdoor Pursuits (e.g., hiking, orienteering, rock climbing)	○	○	○	○	○	○	○	○	○	○	○
Fitness Activities (e.g., strength/weight training, yoga, cycling/biking)	○	○	○	○	○	○	○	○	○	○	○
Aquatic Activities (e.g., swimming, diving)	○	○	○	○	○	○	○	○	○	○	○
Individual Sports (e.g., track & field, self defense, gymnastics)	○	○	○	○	○	○	○	○	○	○	○
Dance/Rhythmic Activities (e.g., creative, line, modern)	○	○	○	○	○	○	○	○	○	○	○
Net/Wall Activities (e.g., tennis, badminton, pickleball)	○	○	○	○	○	○	○	○	○	○	○
Target Activities (e.g., golf, bowling, archery)	○	○	○	○	○	○	○	○	○	○	○

16. PEES-LPA question 10 is clear and concise.

☐ Strongly Disagree ☐ Disagree ☐ Neutral ☐ Agree ☐ Strongly Agree

17. **<u>Identifying and defining</u>** critical skill elements are essential skills for instruction of lifetime physical activities.

☐ Strongly Disagree ☐ Disagree ☐ Neutral ☐ Agree ☐ Strongly Agree

18. Overall, this question is valuable in measuring the construct: *self-efficacy perceptions of PE teachers towards the instruction of lifetime physical activities*.

☐ Strongly Disagree ☐ Disagree ☐ Neutral ☐ Agree ☐ Strongly Agree

19. (Optional) Please provide any overall feedback you may have about question 10.

11. How confident are you in your ability to PRESENT (describe & demonstrate) critical skill elements in the following activity categories?

	0– No Confidence	1	2	3	4	5– Moderately Confident	6	7	8	9	10 – Very Confident
Outdoor Pursuits (e.g., hiking, orienteering, rock climbing)	○	○	○	○	○	○	○	○	○	○	○
Fitness Activities (e.g., strength/weight training, yoga, cycling/biking)	○	○	○	○	○	○	○	○	○	○	○
Aquatic Activities (e.g., swimming, diving)	○	○	○	○	○	○	○	○	○	○	○
Individual Sports (e.g., track & field, self defense, gymnastics)	○	○	○	○	○	○	○	○	○	○	○
Dance/Rhythmic Activities (e.g., creative, line, modern)	○	○	○	○	○	○	○	○	○	○	○
Net/Wall Activities (e.g., tennis, badminton, pickleball)	○	○	○	○	○	○	○	○	○	○	○
Target Activities (e.g., golf, bowling, archery)	○	○	○	○	○	○	○	○	○	○	○

20. PEES-LPA question 11 is clear and concise.

☐ Strongly Disagree ☐ Disagree ☐ Neutral ☐ Agree ☐ Strongly Agree

21. **Presentation** of critical skill elements is an essential skill for instruction of lifetime physical activities.

☐ Strongly Disagree ☐ Disagree ☐ Neutral ☐ Agree ☐ Strongly Agree

22. Overall, this question is valuable in measuring the construct: *self-efficacy perceptions of PE teachers towards the instruction of lifetime physical activities*.

☐ Strongly Disagree ☐ Disagree ☐ Neutral ☐ Agree ☐ Strongly Agree

23. (Optional) Please provide any overall feedback you may have about question 11.

12. How confident are you in your ability to <u>ACCURATELY ASSESS STUDENT SKILL ABILITY</u> of critical skill elements in the following activity categories?

	0– No Confidence	1	2	3	4	5– Moderately Confident	6	7	8	9	10 – Very Confident
Outdoor Pursuits (e.g., hiking, orienteering, rock climbing)	○	○	○	○	○	○	○	○	○	○	○
Fitness Activities (e.g., strength/weight training, yoga, cycling/biking)	○	○	○	○	○	○	○	○	○	○	○
Aquatic Activities (e.g., swimming, diving)	○	○	○	○	○	○	○	○	○	○	○
Individual Sports (e.g., track & field, self defense, gymnastics)	○	○	○	○	○	○	○	○	○	○	○
Dance/Rhythmic Activities (e.g., creative, line, modern)	○	○	○	○	○	○	○	○	○	○	○
Net/Wall Activities (e.g., tennis, badminton, pickleball)	○	○	○	○	○	○	○	○	○	○	○
Target Activities (e.g., golf, bowling, archery)	○	○	○	○	○	○	○	○	○	○	○

24. PEES-LPA question 12 is clear and concise.

☐ Strongly Disagree ☐ Disagree ☐ Neutral ☐ Agree ☐ Strongly Agree

25. <u>**Accurately assessing student skill ability**</u> of critical skill elements is an essential skill for instruction of lifetime physical activities.

☐ Strongly Disagree ☐ Disagree ☐ Neutral ☐ Agree ☐ Strongly Agree

26. Overall, this question is valuable in measuring the construct: *self-efficacy perceptions of PE teachers towards the instruction of lifetime physical activities*.

☐ Strongly Disagree ☐ Disagree ☐ Neutral ☐ Agree ☐ Strongly Agree

27. (Optional) Please provide any overall feedback you may have about question 12.

13. How confident are you in your ability to <u>PROVIDE SKILL-SPECIFIC FEEDBACK</u> of critical skill elements in the following activity categories?

	0– No Confidence	1	2	3	4	5– Moderately Confident	6	7	8	9	10 – Very Confident
Outdoor Pursuits (e.g., hiking, orienteering, rock climbing)	○	○	○	○	○	○	○	○	○	○	○
Fitness Activities (e.g., strength/weight training, yoga, cycling/biking)	○	○	○	○	○	○	○	○	○	○	○
Aquatic Activities (e.g., swimming, diving)	○	○	○	○	○	○	○	○	○	○	○
Individual Sports (e.g., track & field, self defense, gymnastics)	○	○	○	○	○	○	○	○	○	○	○
Dance/Rhythmic Activities (e.g., creative, line, modern)	○	○	○	○	○	○	○	○	○	○	○
Net/Wall Activities (e.g., tennis, badminton, pickleball)	○	○	○	○	○	○	○	○	○	○	○
Target Activities (e.g., golf, bowling, archery)	○	○	○	○	○	○	○	○	○	○	○

28. PEES-LPA question 13 is clear and concise.

☐ Strongly Disagree ☐ Disagree ☐ Neutral ☐ Agree ☐ Strongly Agree

29. <u>**Providing skill-specific feedback**</u> for critical skill elements is an essential skill for instruction of lifetime physical activities.

☐ Strongly Disagree ☐ Disagree ☐ Neutral ☐ Agree ☐ Strongly Agree

30. Overall, this question is valuable in measuring the construct: *self-efficacy perceptions of PE teachers towards the instruction of lifetime physical activities*.

☐ Strongly Disagree ☐ Disagree ☐ Neutral ☐ Agree ☐ Strongly Agree

31. (Optional) Please provide any overall feedback you may have about question 13.

14. How confident are you in your ability to ADAPT INSTRUCTION to meet the needs of both high and low level learners typically found in general physical education classes in the following activity categories?

	0– No Confidence	1	2	3	4	5– Moderately Confident	6	7	8	9	10 – Very Confident
Outdoor Pursuits (e.g., hiking, orienteering, rock climbing)	○	○	○	○	○	○	○	○	○	○	○
Fitness Activities (e.g., strength/weight training, yoga, cycling/biking)	○	○	○	○	○	○	○	○	○	○	○
Aquatic Activities (e.g., swimming, diving)	○	○	○	○	○	○	○	○	○	○	○
Individual Sports (e.g., track & field, self defense, gymnastics)	○	○	○	○	○	○	○	○	○	○	○
Dance/Rhythmic Activities (e.g., creative, line, modern)	○	○	○	○	○	○	○	○	○	○	○
Net/Wall Activities (e.g., tennis, badminton, pickleball)	○	○	○	○	○	○	○	○	○	○	○
Target Activities (e.g., golf, bowling, archery)	○	○	○	○	○	○	○	○	○	○	○

32. PEES-LPA question 14 is clear and concise.

☐ Strongly Disagree ☐ Disagree ☐ Neutral ☐ Agree ☐ Strongly Agree

33. **Adapting instruction** for students typically found in general physical education class is an essential skill for instruction of lifetime physical activities.

☐ Strongly Disagree ☐ Disagree ☐ Neutral ☐ Agree ☐ Strongly Agree

34. Overall, this question is valuable in measuring the construct: *self-efficacy perceptions of PE teachers towards the instruction of lifetime physical activities*.

☐ Strongly Disagree ☐ Disagree ☐ Neutral ☐ Agree ☐ Strongly Agree

35. (Optional) Please provide any overall feedback you may have about question 14.

15. How confident are you in your ability to <u>ADDRESS THE NEEDS OF STUDENTS WITH DISABILITIES (physical, sensory, or intellectual)</u> and are included in your physical education classes in the following activity categories?

	0– No Confidence	1	2	3	4	5– Moderately Confident	6	7	8	9	10 – Very Confident
Outdoor Pursuits (e.g., hiking, orienteering, rock climbing)											
Fitness Activities (e.g., strength/weight training, yoga, cycling/biking)											
Aquatic Activities (e.g., swimming, diving)											
Individual Sports (e.g., track & field, self defense, gymnastics)											
Dance/Rhythmic Activities (e.g., creative, line, modern)											
Net/Wall Activities (e.g., tennis, badminton, pickleball)											
Target Activities (e.g., golf, bowling, archery)											

36. PEES-LPA question 15 is clear and concise.

☐ Strongly Disagree ☐ Disagree ☐ Neutral ☐ Agree ☐ Strongly Agree

37. **<u>Addressing the needs of students with disabilities</u>** who are included in physical education classes is an essential skill for instruction of lifetime physical activities.

☐ Strongly Disagree ☐ Disagree ☐ Neutral ☐ Agree ☐ Strongly Agree

38. Overall, this question is valuable in measuring the construct: *self-efficacy perceptions of PE teachers towards the instruction of lifetime physical activities*.

☐ Strongly Disagree ☐ Disagree ☐ Neutral ☐ Agree ☐ Strongly Agree

39. (Optional) Please provide any overall feedback you may have about question 15.

CONCLUSION

(Optional) Please provide any overall feedback you may have about the PEES-LPA instrument:

Thank you for your time and feedback. Please save the file (as a word or PDF document) and send the completed form to Kason O'Neil at kmo7rh@virginia.edu

Appendix D

Physical Educator Efficacy Scale Towards Teaching Lifetime Physical Activities (PEES-LPA)

Dear Physical Educator,

Please read this consent agreement carefully before you decide to participate in this study.

You are invited to participate in a brief research study with the purpose of collecting information on lifetime physical activities in secondary physical education. In this study, you will be given a list of lifetime physical activities (drawn from a 2013 AAHPERD publication), and asked to answer survey questions referring to confidence levels and personal experience towards instruction of specific lifetime physical activity categories.

This survey will take approximately 5-10 minutes to complete.

Your participation in this study is completely voluntary. There are no direct benefits to you for participating in this research study. Additionally, there are no foreseeable risks associated with this project. However, if you feel uncomfortable answering any questions, feel free to skip on to the next question. Your survey responses will be strictly confidential and data from this research will be reported only in the aggregate. Your information will be coded and will remain confidential. If you have questions at any time about the survey or the procedures, you may contact:

Kason O'Neil
University of Virginia, Charlottesville, VA 22903
518-368-0708
kmo7rh@virginia.edu

If you have questions about your rights in the study, contact:
Tonya R. Moon, PhD.D.Chair,
Institutional Review Board for the Social and Behavioral Sciences
One Morton Dr Suite 500University of Virginia, P.O. Box 800392
Charlottesville, VA 22908-0392
Telephone: (434) 924-5999
Email: irbsbshelp@virginia.edu

Thank you very much for your time and cooperation.

Consent Statement:
I agree to participate in the research study described above.

Name______________________________

Signature__________________________

Date______________________________

Demographic Questions

Participant Email: ____

Which grade(s) do you currently teach secondary physical education?

- ❍ Middle School (6-8)
- ❍ High School (9-12)
- ❍ Both MS & HS
- ❍ None of the above

In what state are you currently teaching as a certified physical educator? ________

How many years of experience do you have as a certified physical education teacher? __________

What is your highest educational degree?

- ❍ Bachelors
- ❍ Masters
- ❍ Doctorate

Gender?

- ❍ Male
- ❍ Female

FOR THIS STUDY…

Lifetime Physical Activities are defined as being any physical activity that may be carried over into adulthood due to being accomplished by one or two people, and requiring little structure, organization, and equipment. The purpose of the following survey will be to measure your perceptions and confidence levels towards teaching Lifetime Physical Activities in secondary physical education.

Note: For all of the following questions in this survey, responses will refer to each activity category **as a whole,** and not to specific activity examples. For example, you will be asked to answer perception questions in the category of 'fitness activities'. Your answer will represent an overall analysis of a combination of 'fitness activities' such as strength training, yoga, cycling, etc. combined, and not towards selected individual activities.

Perceptions of Experience with Lifetime Physical Activities

1. Please rate how you would classify your ***personal skill-ability*** within each activity category.

	No Experience	Novice	Advanced Beginner	Competent	Proficient	Expert
Outdoor Pursuits (e.g., hiking, orienteering, rock climbing)	❍	❍	❍	❍	❍	❍
Fitness Activities (e.g., strength/weight training, yoga, cycling/biking)	❍	❍	❍	❍	❍	❍
Aquatic Activities (e.g., swimming, diving)	❍	❍	❍	❍	❍	❍
Individual Sports (e.g., track & field, self defense, gymnastics)	❍	❍	❍	❍	❍	❍
Dance/Rhythmic Activities (e.g., creative, line, modern)	❍	❍	❍	❍	❍	❍
Net/Wall Activities (e.g., tennis, badminton, pickleball)	❍	❍	❍	❍	❍	❍
Target Activities (e.g., golf, bowling, archery)	❍	❍	❍	❍	❍	❍

2. On average, how frequently have you participated in the following activity categories (competitively or recreationally) in the past YEAR?

	Never	Only a few times	Once a month	Weekly	More than once a week
Outdoor Pursuits (e.g., hiking, orienteering, rock climbing)	❍	❍	❍	❍	❍
Fitness Activities (e.g., strength/weight training, yoga, cycling/biking)	❍	❍	❍	❍	❍
Aquatic Activities (e.g., swimming, diving)	❍	❍	❍	❍	❍
Individual Sports (e.g., track & field, self defense, gymnastics)	❍	❍	❍	❍	❍
Dance/Rhythmic Activities (e.g., creative, line, modern)	❍	❍	❍	❍	❍
Net/Wall Activities (e.g., tennis, badminton, pickleball)	❍	❍	❍	❍	❍
Target Activities (e.g., golf, bowling, archery)	❍	❍	❍	❍	❍

3. Please rate how you would classify your competence to teach a ***variety of activities*** within each activity category.

	Novice	Advanced Beginner	Competent	Proficient	Expert
Outdoor Pursuits (e.g., hiking, orienteering, rock climbing)	❍	❍	❍	❍	❍
Fitness Activities (e.g., strength/weight training, yoga, cycling/biking)	❍	❍	❍	❍	❍
Aquatic Activities (e.g., swimming, diving)	❍	❍	❍	❍	❍
Individual Sports (e.g., track & field, self defense, gymnastics)	❍	❍	❍	❍	❍
Dance/Rhythmic Activities (e.g., creative, line, modern)	❍	❍	❍	❍	❍
Net/Wall Activities (e.g., tennis, badminton, pickleball)	❍	❍	❍	❍	❍
Target Activities (e.g., golf, bowling, archery)	❍	❍	❍	❍	❍

Confidence Towards Instruction of Lifetime Physical Activities

4. How confident are you in your ability to ***IDENTIFY AND DEFINE*** critical skill elements in the following activity categories?

	No Confidence					Moderately Confident					Very Confident
	0	1	2	3	4	5	6	7	8	9	10
Outdoor Pursuits (e.g., hiking, orienteering, rock climbing)	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍
Fitness Activities (e.g., strength/weight training, yoga, cycling/biking)	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍
Aquatic Activities (e.g., swimming, diving)	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍
Individual Sports (e.g., track & field, self defense, gymnastics)	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍
Dance/Rhythmic Activities (e.g., creative, line, modern)	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍
Net/Wall Activities (e.g., tennis, badminton, pickle-ball)	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍
Target Activities (e.g., golf, bowling, archery)	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍

5. How confident are you in your ability to ***PRESENT*** (describe & demonstrate) critical skill elements in the following activity categories?

	No Confidence					Moderately Confident					Very Confident
	0	1	2	3	4	5	6	7	8	9	10
Outdoor Pursuits (e.g., hiking, orienteering, rock climbing)	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍
Fitness Activities (e.g., strength/weight training, yoga, cycling/biking)	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍
Aquatic Activities (e.g., swimming, diving)	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍
Individual Sports (e.g., track & field, self defense, gymnastics)	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍
Dance/Rhythmic Activities (e.g., creative, line, modern)	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍
Net/Wall Activities (e.g., tennis, badminton, pickle-ball)	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍
Target Activities (e.g., golf, bowling, archery)	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍

6. How confident are you in your ability to ***ACCURATELY ASSESS STUDENT SKILL-ABILITY*** of critical skill elements in the following activity categories?

	No Confidence					Moderately Confident					Very Confident
	0	1	2	3	4	5	6	7	8	9	10
Outdoor Pursuits (e.g., hiking, orienteering, rock climbing)	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍
Fitness Activities (e.g., strength/weight training, yoga, cycling/biking)	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍
Aquatic Activities (e.g., swimming, diving)	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍
Individual Sports (e.g., track & field, self defense, gymnastics)	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍
Dance/Rhythmic Activities (e.g., creative, line, modern)	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍
Net/Wall Activities (e.g., tennis, badminton, pickle-ball)	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍
Target Activities (e.g., golf, bowling, archery)	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍

7. How confident are you in your ability to ***PROVIDE SKILL-SPECIFIC FEEDBACK*** of critical skill elements in the following activity categories?

	No Confidence					Moderately Confident					Very Confident
	0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Outdoor Pursuits (e.g., hiking, orienteering, rock climbing)	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍
Fitness Activities (e.g., strength/weight training, yoga, cycling/biking)	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍
Aquatic Activities (e.g., swimming, diving)	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍
Individual Sports (e.g., track & field, self defense, gymnastics)	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍
Dance/Rhythmic Activities (e.g., creative, line, modern)	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍
Net/Wall Activities (e.g., tennis, badminton, pickle-ball)	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍
Target Activities (e.g., golf, bowling, archery)	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍

8. How confident are you in your ability to ***DIFFERENTIATE INSTRUCTION*** to meet the needs of both high and low level learners typically found in general physical education classes in the following activity categories?

	No Confidence					Moderately Confident					Very Confident
	0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Outdoor Pursuits (e.g., hiking, orienteering, rock climbing)	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍
Fitness Activities (e.g., strength/weight training, yoga, cycling/biking)	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍
Aquatic Activities (e.g., swimming, diving)	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍
Individual Sports (e.g., track & field, self defense, gymnastics)	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍
Dance/Rhythmic Activities (e.g., creative, line, modern)	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍
Net/Wall Activities (e.g., tennis, badminton, pickle-ball)	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍
Target Activities (e.g., golf, bowling, archery)	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍

9. How confident are you in your ability to ***ADDRESS THE NEEDS OF STUDENTS WITH DISABILITIES*** (physical, sensory, or intellectual) who may be included in your physical education classes in the following activity categories?

	No Confidence					Moderately Confident					Very Confident
	0	1	2	3	4	5	6	7	8	9	10
Outdoor Pursuits (e.g., hiking, orienteering, rock climbing)	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍
Fitness Activities (e.g., strength/weight training, yoga, cycling/biking)	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍
Aquatic Activities (e.g., swimming, diving)	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍
Individual Sports (e.g., track & field, self defense, gymnastics)	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍
Dance/Rhythmic Activities (e.g., creative, line, modern)	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍
Net/Wall Activities (e.g., tennis, badminton, pickle-ball)	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍
Target Activities (e.g., golf, bowling, archery)	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍	❍

Appendix E

Skewness and Kurtosis for PEES-LPA Variables

Skewness and Kurtosis of PEES-LPA Variables

		N	Skewness		Kurtosis	
		Statistic	Statistic	Std. Error	Statistic	Std. Error
1	SkillAbility_Outdoor	182	-0.218	0.18	-0.941	0.358
2	SkillAbility_Fitness	182	-0.702	0.18	0.474	0.358
3	SkillAbility_Aquatic	182	-0.481	0.18	-0.451	0.358
4	SkillAbility_Ind	182	-0.733	0.18	0.284	0.358
5	SkillAbility_Dance	182	-0.334	0.18	-0.511	0.358
6	SkilAbility_Net	182	-0.748	0.18	0.537	0.358
7	SkillAbilty_Target	182	-0.77	0.18	0.342	0.358
8	Trans_Freq_Fit	182	0.676	0.18	-0.687	0.358
9	Freq_Outdooor	182	0.716	0.18	0.378	0.358
10	Freq_Aquatic	182	0.661	0.18	0.205	0.358
11	Freq_Ind	182	0.575	0.18	-0.818	0.358
12	Freq_Dance	182	0.806	0.18	0.03	0.358
13	Freq_Net	182	0.336	0.18	-0.512	0.358
14	Freq_Target	182	0.553	0.18	-0.246	0.358
15	Tran_Variety_Fitness	182	-0.252	0.18	-0.937	0.358
16	Tran_Variety_Net	182	-0.344	0.18	-0.482	0.358
17	Variety_Outdoor	182	-0.263	0.18	-0.997	0.358
18	Variety_Aquatic	182	-0.203	0.18	-0.946	0.358
19	Variety_Ind	182	-0.634	0.18	0.317	0.358
20	Variety_Dance	182	-0.256	0.18	-0.798	0.358
21	Variety_Target	182	-0.917	0.18	1.303	0.358
22	Tran_Identify_Fit	182	0.011	0.18	-0.918	0.358
23	Tran_Identify_Net	182	-0.122	0.18	-0.889	0.358
24	Tran_Identify_Target	182	-0.138	0.18	-0.803	0.358
25	Identify_Outdoor	182	-0.3	0.18	-0.48	0.358
26	Identify_Aquatic	182	-0.395	0.18	-0.956	0.358
27	Identify_Ind	182	-0.82	0.18	0.331	0.358
28	Identify_Dance	182	-0.291	0.18	-0.94	0.358
29	Tran_Present_Fitness	182	0.095	0.18	-0.998	0.358
30	Trans_Present_Net	182	0.001	0.18	-1.047	0.358
31	Present_Outdoor	182	-0.364	0.18	-0.556	0.358
32	Present_Aquatic	182	-0.426	0.18	-1.001	0.358
33	Present_Ind	182	-0.85	0.18	0.366	0.358
34	Present_Dance	182	-0.385	0.18	-0.981	0.358
35	Present_Target	182	-0.986	0.18	0.533	0.358
36	Tran_Assess_Fitness	182	0.225	0.18	-1.144	0.358
37	Tran_Assess_Ind	182	-0.11	0.18	-0.854	0.358
38	Tran_Assess_Net	182	0.248	0.18	-0.984	0.358
39	Tran_Assess_Target	182	0.153	0.18	-1.122	0.358
40	Assess_Outdoor	182	-0.382	0.18	-0.74	0.358

41	Assess_Aquatic	182	-0.585	0.18	-0.81	0.358
42	Assess_Dance	182	-0.552	0.18	-0.828	0.358
43	Tran_FB_Fitness	182	0.178	0.18	-1.079	0.358
44	Tran_FB_Ind	182	-0.136	0.18	-0.742	0.358
45	Tran_FB_Net	182	0.162	0.18	-0.894	0.358
46	Tran_FB_Target	182	0.01	0.18	-1.048	0.358
47	Feedback_Outdoor	182	-0.362	0.18	-0.778	0.358
48	Feedback_Aquatic	182	-0.519	0.18	-0.852	0.358
49	Feedback_Dace	182	-0.507	0.18	-0.878	0.358
50	Trans_Diff_Fitness	182	0.227	0.18	-1.167	0.358
51	Trans_Diff_Ind	182	-0.149	0.18	-1.033	0.358
52	Trans_Diff_Target	182	0.127	0.18	-1.121	0.358
53	Diff_Outdoor	182	-0.38	0.18	-0.824	0.358
54	Diff_Aquatic	182	-0.549	0.18	-0.801	0.358
55	Diff_Dance	182	-0.436	0.18	-0.972	0.358
56	Tran_Diff_Net	182	0.232	0.18	-0.941	0.358
57	SSN_Outdoor	182	-0.33	0.18	-0.941	0.358
58	SSN_Aquatic	182	-0.275	0.18	-1.114	0.358
59	SSN_Ind	182	-0.883	0.18	-0.041	0.358
60	SSN_Dance	182	-0.441	0.18	-0.771	0.358
61	Tran_SSN_Fit	182	0.028	0.18	-1.205	0.358
62	Tran_SSN_Net	182	-0.03	0.18	-1.192	0.358
63	Tran_SSN_Target	182	-0.062	0.18	-1.152	0.358
	Valid N (listwise)	182				

Appendix F

Inter-Item Correlation Matrix of PEES-LPA Variables

	SkillAbility_Outdoor	SkillAbility_Fitness	SkillAbility_Aquatic	SkillAbility_Ind	SkillAbility_Dance	SkilAbility_Net	SkillAbilty_Target	Freq_Outdoor	Freq_Aquatic	Freq_Ind	Freq_Dance	Freq_Net	Freq_Fitness	Freq_Target	Variety_Outdoor	Variety_Aquatic	Variety_Ind	Variety_Dance	Variety_Target	Variety_Net
SkillAbility_Outdoor	1	0.285	0.358	0.175	0.171	0.167	0.181	0.566	0.219	0.223	-0.1	0.026	0.065	0.089	0.778	0.295	0.099	-0.007	0.088	0.158
SkillAbility_Fitness	0.285	1	0.278	0.364	0.25	0.384	0.191	0.167	0.039	0.229	0.094	-0.019	0.407	-0.146	0.299	0.27	0.359	0.176	0.246	0.389
SkillAbility_Aquatic	0.358	0.278	1	0.418	0.324	0.237	0.113	0.176	0.499	0.237	0.052	0.032	-0.003	-0.079	0.299	0.841	0.333	0.134	-0.004	0.179
SkillAbility_Ind	0.175	0.364	0.418	1	0.368	0.261	0.173	0.073	0.249	0.433	0.199	0.057	0.118	-0.12	0.26	0.442	0.759	0.264	0.214	0.197
SkillAbility_Dance	0.171	0.25	0.324	0.368	1	0.208	0.09	0.023	0.125	0.221	0.65	-0.011	0.121	-0.039	0.211	0.351	0.334	0.835	0.019	0.15
SkilAbility_Net	0.167	0.384	0.237	0.261	0.208	1	0.583	0.012	0.065	0.109	0.013	0.304	0.047	0.321	0.214	0.263	0.276	0.139	0.443	0.768
SkillAbilty_Target	0.181	0.191	0.113	0.173	0.09	0.583	1	-0.012	-0.001	0.112	-0.14	0.111	0.056	0.425	0.227	0.18	0.162	0.071	0.7	0.501
Freq_Outdooor	0.566	0.167	0.176	0.073	0.023	0.012	-0.012	1	0.123	0.151	-0.048	0.108	0.112	0.094	0.543	0.141	-0.021	-0.099	-0.134	0.043
Freq_Aquatic	0.219	0.039	0.499	0.249	0.125	0.065	-0.001	0.123	1	0.373	0.093	0.11	0.095	0.122	0.173	0.52	0.249	0.028	0.017	0.067
Freq_Ind	0.223	0.229	0.237	0.433	0.221	0.109	0.112	0.151	0.373	1	0.19	0.122	0.315	0.048	0.232	0.282	0.483	0.2	0.114	0.101
Freq_Dance	-0.1	0.094	0.052	0.199	0.65	0.013	-0.14	-0.048	0.093	0.19	1	0.11	0.197	-0.007	-0.029	0.092	0.208	0.693	-0.079	-0.008
Freq_Net	0.026	-0.019	0.032	0.057	-0.011	0.304	0.111	0.108	0.11	0.122	0.11	1	0.085	0.471	0.042	-0.023	0.059	0.019	0.074	0.322
Freq_Fitness	0.065	0.407	-0.003	0.118	0.121	0.047	0.056	0.112	0.095	0.315	0.197	0.085	1	0.056	0.024	0.001	0.192	0.142	0.094	-0.024
Freq_Target	0.089	-0.146	-0.079	-0.12	-0.039	0.321	0.425	0.094	0.122	0.048	-0.007	0.471	0.056	1	0.081	-0.008	-0.064	0.02	0.354	0.273
Variety_Outdoor	0.778	0.299	0.299	0.26	0.211	0.214	0.227	0.543	0.173	0.232	-0.029	0.042	0.024	0.081	1	0.314	0.228	0.123	0.157	0.256
Variety_Aquatic	0.295	0.27	0.841	0.442	0.351	0.263	0.18	0.141	0.52	0.282	0.092	-0.023	0.001	-0.008	0.314	1	0.458	0.226	0.105	0.246
Variety_Ind	0.099	0.359	0.333	0.759	0.334	0.276	0.162	-0.021	0.249	0.483	0.208	0.059	0.192	-0.064	0.228	0.458	1	0.314	0.299	0.316
Variety_Dance	-0.007	0.176	0.134	0.264	0.835	0.139	0.071	-0.099	0.028	0.2	0.693	0.019	0.142	0.02	0.123	0.226	0.314	1	0.09	0.177
Variety_Target	0.088	0.246	-0.004	0.214	0.019	0.443	0.7	-0.134	0.017	0.114	-0.079	0.074	0.094	0.354	0.157	0.105	0.299	0.09	1	0.587
Variety_Net	0.158	0.389	0.179	0.197	0.15	0.768	0.501	0.043	0.067	0.101	-0.008	0.322	-0.024	0.273	0.256	0.246	0.316	0.177	0.587	1
Variety_Fitness	0.19	0.776	0.205	0.326	0.228	0.384	0.2	0.098	0.067	0.276	0.089	0.057	0.41	-0.031	0.265	0.274	0.407	0.2	0.283	0.455
Identify_Outdoor	0.644	0.254	0.28	0.221	0.179	0.052	0.158	0.456	0.12	0.196	-0.003	-0.078	0.084	0.032	0.728	0.245	0.198	0.112	0.097	0.073
Identify_Target	0.044	0.115	0.083	0.213	0.094	0.356	0.583	-0.023	0.003	0.031	-0.05	0.053	-0.012	0.434	0.088	0.18	0.217	0.111	0.577	0.368
Identify_Net	0.001	0.252	0.189	0.211	0.169	0.577	0.436	0.016	0.057	0.04	0.026	0.29	-0.055	0.285	0.095	0.227	0.237	0.166	0.387	0.639
Identify_Fitness	0.135	0.649	0.21	0.333	0.289	0.208	0.057	0.147	0.05	0.154	0.16	-0.056	0.416	-0.046	0.199	0.207	0.349	0.239	0.108	0.224
Identify_Aquatic	0.23	0.155	0.783	0.351	0.374	0.212	0.174	0.105	0.457	0.26	0.203	0.019	0.07	0.054	0.218	0.843	0.373	0.267	0.082	0.169
Identify_Ind	0.087	0.311	0.334	0.702	0.36	0.223	0.172	-0.01	0.18	0.456	0.197	0.021	0.1	-0.055	0.189	0.441	0.759	0.349	0.195	0.245
Identify_Dance	-0.062	0.123	0.128	0.255	0.787	0.044	0.02	-0.067	0.033	0.14	0.713	-0.019	0.165	-0.078	0.059	0.19	0.301	0.843	0.027	0.067
Present_Outdoor	0.708	0.278	0.272	0.246	0.146	0.108	0.156	0.529	0.113	0.23	-0.053	-0.046	0.119	0.051	0.76	0.234	0.18	0.052	0.092	0.112
Present_Aquatic	0.241	0.166	0.785	0.377	0.332	0.205	0.134	0.125	0.46	0.265	0.205	0.005	0.089	-0.008	0.222	0.836	0.373	0.229	0.026	0.14
Present_Ind	0.125	0.272	0.385	0.673	0.387	0.187	0.102	0.059	0.223	0.443	0.194	0.003	0.176	-0.028	0.203	0.486	0.739	0.324	0.15	0.2
Present_Dance	-0.057	0.084	0.101	0.232	0.75	0.028	0.001	-0.059	0.035	0.101	0.703	-0.016	0.164	-0.079	0.034	0.15	0.255	0.81	-0.005	0.01
Present_Target	0.133	0.115	0.054	0.228	0.104	0.393	0.58	-0.015	-0.057	-0.004	-0.085	0.06	-0.003	0.367	0.154	0.127	0.186	0.115	0.585	0.432
Present_Fitness	0.145	0.59	0.186	0.341	0.295	0.244	0.108	0.141	-0.017	0.153	0.185	-0.025	0.426	0.018	0.169	0.164	0.364	0.259	0.121	0.224
Present_Net	0.053	0.227	0.168	0.262	0.236	0.615	0.484	0.006	0.057	0.091	0.056	0.287	0.012	0.286	0.101	0.227	0.279	0.229	0.464	0.67
Assess_Outdoor	0.675	0.286	0.281	0.27	0.184	0.171	0.204	0.497	0.157	0.238	-0.016	0.028	0.05	0.09	0.764	0.258	0.228	0.079	0.176	0.197
Assess_Aquatic	0.259	0.234	0.762	0.379	0.385	0.252	0.19	0.124	0.422	0.271	0.191	0.014	0.08	0.01	0.27	0.816	0.395	0.267	0.104	0.233

	Variety_Fitness	Identify_Outdoor	Identify_Target	Identify_Net	Identify_Fitness	Identify_Aquatic	Identify_Ind	Identify_Dance	Present_Outdoor	Present_Aquatic	Present_Ind	Present_Dance	Present_Target	Present_Fitness	Present_Net	Assess_Outdoor	Assess_Aquatic	Assess_Dance	Assess_Fitness	Assess_Ind	Assess_Net	Assess_Target
SkillAbility_Outdoor	0.19	0.644	0.044	0.001	0.135	0.23	0.087	-0.062	0.708	0.241	0.125	-0.057	0.133	0.145	0.053	0.675	0.259	-0.007	0.125	0.094	0.054	0.085
SkillAbility_Fitness	0.776	0.254	0.115	0.252	0.649	0.155	0.311	0.123	0.278	0.166	0.272	0.084	0.115	0.59	0.227	0.286	0.234	0.073	0.534	0.252	0.222	0.143
SkillAbility_Aquatic	0.205	0.28	0.083	0.189	0.21	0.783	0.334	0.128	0.272	0.785	0.385	0.101	0.054	0.186	0.168	0.281	0.762	0.16	0.17	0.323	0.128	0.082
SkillAbility_Ind	0.326	0.221	0.213	0.211	0.333	0.351	0.702	0.255	0.246	0.377	0.673	0.232	0.228	0.341	0.262	0.27	0.379	0.282	0.335	0.655	0.223	0.273
SkillAbility_Dance	0.228	0.179	0.094	0.169	0.289	0.374	0.36	0.787	0.146	0.332	0.387	0.75	0.104	0.295	0.236	0.184	0.385	0.76	0.228	0.368	0.196	0.135
SkilAbility_Net	0.384	0.052	0.356	0.577	0.208	0.212	0.223	0.044	0.108	0.205	0.187	0.028	0.393	0.244	0.615	0.171	0.252	0.069	0.229	0.215	0.617	0.43
SkillAbilty_Target	0.2	0.158	0.583	0.436	0.057	0.174	0.172	0.02	0.156	0.134	0.102	0.001	0.58	0.108	0.484	0.204	0.19	-0.002	0.092	0.12	0.409	0.495
Freq_Outdooor	0.098	0.456	-0.023	0.016	0.147	0.105	-0.01	-0.067	0.529	0.125	0.059	-0.059	-0.015	0.141	0.006	0.497	0.124	-0.048	0.133	0.02	0.022	0.012
Freq_Aquatic	0.067	0.12	0.003	0.057	0.05	0.457	0.18	0.033	0.113	0.46	0.223	0.035	-0.057	-0.017	0.057	0.157	0.422	0.067	0.038	0.186	0.073	0.017
Freq_Ind	0.276	0.196	0.031	0.04	0.154	0.26	0.456	0.14	0.23	0.265	0.443	0.101	-0.004	0.153	0.091	0.238	0.271	0.166	0.146	0.429	0.105	0.069
Freq_Dance	0.089	-0.003	-0.05	0.026	0.16	0.203	0.197	0.713	-0.053	0.205	0.194	0.703	-0.085	0.185	0.056	-0.016	0.191	0.635	0.145	0.196	0.084	0.023
Freq_Net	0.057	-0.078	0.053	0.29	-0.056	0.019	0.021	-0.019	-0.046	0.005	0.003	-0.016	0.06	-0.025	0.287	0.028	0.014	0.03	-0.065	0.047	0.306	0.074
Freq_Fitness	0.41	0.084	-0.012	-0.055	0.416	0.07	0.1	0.165	0.119	0.089	0.176	0.164	-0.003	0.426	0.012	0.05	0.08	0.133	0.261	0.124	-0.018	0.015
Freq_Target	-0.031	0.032	0.434	0.285	-0.046	0.054	-0.055	-0.078	0.051	-0.008	-0.028	-0.079	0.367	0.018	0.286	0.09	0.01	-0.073	0.004	-0.025	0.258	0.354
Variety_Outdoor	0.265	0.728	0.088	0.095	0.199	0.218	0.189	0.059	0.76	0.222	0.203	0.034	0.154	0.169	0.101	0.764	0.27	0.095	0.182	0.176	0.16	0.149
Variety_Aquatic	0.274	0.245	0.18	0.227	0.207	0.843	0.441	0.19	0.234	0.836	0.486	0.15	0.127	0.164	0.227	0.258	0.816	0.237	0.224	0.444	0.233	0.215
Variety_Ind	0.407	0.198	0.217	0.237	0.349	0.373	0.759	0.301	0.18	0.373	0.739	0.255	0.186	0.364	0.279	0.228	0.395	0.319	0.366	0.753	0.27	0.258
Variety_Dance	0.2	0.112	0.111	0.166	0.239	0.267	0.349	0.843	0.052	0.229	0.324	0.81	0.115	0.259	0.229	0.079	0.267	0.794	0.182	0.329	0.198	0.169
Variety_Target	0.283	0.097	0.577	0.387	0.108	0.082	0.195	0.027	0.092	0.026	0.15	-0.005	0.585	0.121	0.464	0.176	0.104	0.027	0.165	0.186	0.446	0.601
Variety_Net	0.455	0.073	0.368	0.639	0.224	0.169	0.245	0.067	0.112	0.14	0.2	0.01	0.432	0.224	0.67	0.197	0.233	0.085	0.288	0.254	0.669	0.492
Variety_Fitness	1	0.191	0.143	0.297	0.683	0.149	0.326	0.151	0.211	0.155	0.311	0.116	0.131	0.63	0.286	0.249	0.22	0.12	0.613	0.309	0.313	0.221
Identify_Outdoor	0.191	1	0.186	0.077	0.333	0.313	0.24	0.166	**0.883**	0.273	0.184	0.115	0.162	0.3	0	0.819	0.296	0.176	0.221	0.176	0.051	0.174
Identify_Target	0.143	0.186	1	0.635	0.214	0.271	0.375	0.118	0.134	0.184	0.3	0.11	0.831	0.263	0.574	0.164	0.24	0.085	0.223	0.298	0.516	0.785
Identify_Net	0.297	0.077	0.635	1	0.302	0.273	0.384	0.191	0.041	0.201	0.315	0.126	0.586	0.317	0.82	0.141	0.288	0.188	0.297	0.328	0.783	0.606
Identify_Fitness	0.683	0.333	0.214	0.302	1	0.27	0.433	0.249	0.297	0.218	0.373	0.196	0.161	0.833	0.217	0.298	0.246	0.213	0.717	0.356	0.225	0.212
Identify_Aquatic	0.149	0.313	0.271	0.273	0.27	1	0.479	0.327	0.248	**0.912**	0.493	0.253	0.188	0.247	0.254	0.247	**0.88**	0.335	0.236	0.442	0.228	0.245
Identify_Ind	0.326	0.24	0.375	0.384	0.433	0.479	1	0.388	0.164	0.408	0.815	0.292	0.28	0.4	0.345	0.192	0.436	0.369	0.366	0.823	0.344	0.348
Identify_Dance	0.151	0.166	0.118	0.191	0.249	0.327	0.388	1	0.072	0.261	0.353	**0.907**	0.099	0.233	0.216	0.077	0.304	**0.894**	0.152	0.339	0.182	0.159
Present_Outdoor	0.211	**0.883**	0.134	0.041	0.297	0.248	0.164	0.072	1	0.309	0.246	0.103	0.215	0.329	0.072	**0.883**	0.321	0.136	0.259	0.188	0.077	0.188
Present_Aquatic	0.155	0.273	0.184	0.201	0.218	**0.912**	0.408	0.261	0.309	1	0.535	0.289	0.17	0.271	0.26	0.284	**0.919**	0.309	0.24	0.447	0.202	0.212
Present_Ind	0.311	0.184	0.3	0.315	0.373	0.493	0.815	0.353	0.246	0.535	1	0.352	0.327	0.47	0.407	0.247	0.519	0.383	0.393	0.842	0.328	0.329
Present_Dance	0.116	0.115	0.11	0.126	0.196	0.253	0.292	**0.907**	0.103	0.289	0.352	1	0.151	0.256	0.239	0.087	0.306	**0.891**	0.177	0.323	0.163	0.158
Present_Target	0.131	0.162	0.831	0.586	0.161	0.188	0.28	0.099	0.215	0.17	0.327	0.151	1	0.315	0.734	0.238	0.235	0.127	0.256	0.29	0.596	0.812
Present_Fitness	0.63	0.3	0.263	0.317	0.833	0.247	0.4	0.233	0.329	0.271	0.47	0.256	0.315	1	0.366	0.285	0.275	0.211	0.782	0.391	0.288	0.279
Present_Net	0.286	0	0.574	0.82	0.217	0.254	0.345	0.216	0.072	0.26	0.407	0.239	0.734	0.366	1	0.138	0.332	0.237	0.319	0.371	0.832	0.636
Assess_Outdoor	0.249	0.819	0.164	0.141	0.298	0.247	0.192	0.077	**0.883**	0.284	0.247	0.087	0.238	0.285	0.138	1	0.364	0.175	0.345	0.262	0.232	0.29
Assess_Aquatic	0.22	0.296	0.24	0.288	0.246	**0.88**	0.436	0.304	0.321	**0.919**	0.519	0.306	0.235	0.275	0.332	0.364	1	0.392	0.334	0.532	0.321	0.326

	Feedback_Aquatic	Feedback_Outdoor	Feedback_Fitness	Feedback_Dance	Feedback_Ind	Feedback_Net	Feedback_Target	Diff_Outdoor	Diff_Aquatic	Diff_Dance	Diff_Fitness	Diff_Ind	Diff_Target	Diff_Net	SSN_Outdoor	SSN_Aquatic	SSN_Ind	SSN_Dance	SSN_Fitness	SSN_Net	SSN_Target
SkillAbility_Outdoor	0.276	0.687	0.149	-0.018	0.14	0.097	0.107	0.643	0.27	0.002	0.155	0.132	0.141	0.126	0.563	0.308	0.166	0.063	0.174	0.107	0.146
SkillAbility_Fitness	0.257	0.304	0.576	0.1	0.317	0.233	0.121	0.314	0.238	0.109	0.577	0.3	0.274	0.291	0.309	0.229	0.321	0.138	0.493	0.296	0.281
SkillAbility_Aquatic	0.764	0.315	0.18	0.126	0.335	0.161	0.107	0.282	0.732	0.158	0.251	0.307	0.125	0.178	0.369	0.69	0.314	0.229	0.252	0.2	0.158
SkillAbility_Ind	0.363	0.293	0.342	0.263	0.676	0.249	0.251	0.268	0.395	0.266	0.313	0.599	0.252	0.234	0.338	0.393	0.531	0.348	0.385	0.318	0.305
SkilAbility_Dance	0.374	0.204	0.242	0.744	0.347	0.224	0.163	0.207	0.408	0.735	0.274	0.41	0.26	0.299	0.245	0.381	0.329	0.67	0.211	0.215	0.2
SkilAbility_Net	0.289	0.194	0.256	0.087	0.259	0.645	0.471	0.145	0.248	0.058	0.228	0.209	0.429	0.558	0.1	0.222	0.228	0.105	0.229	0.421	0.334
SkillAbilty_Target	0.226	0.227	0.102	0	0.194	0.449	0.548	0.15	0.171	0.017	0.079	0.082	0.428	0.37	0.129	0.203	0.104	0.081	0.117	0.295	0.337
Freq_Outdooor	0.124	0.486	0.15	-0.075	0.03	0.033	-0.02	0.498	0.155	-0.048	0.138	0.09	0.028	0.089	0.417	0.116	0.069	0.006	0.152	0.052	0.034
Freq_Aquatic	0.404	0.158	0.003	0.024	0.145	0.071	-0.005	0.179	0.415	0.055	0.09	0.196	0.001	0.083	0.226	0.422	0.194	0.127	0.113	0.112	0.067
Freq_Ind	0.261	0.248	0.16	0.143	0.399	0.126	0.04	0.257	0.31	0.167	0.179	0.44	0.074	0.159	0.287	0.329	0.371	0.195	0.215	0.226	0.167
Freq_Dance	0.182	-0.005	0.161	0.648	0.189	0.095	0.041	0.042	0.18	0.613	0.114	0.223	0.113	0.149	0.083	0.166	0.156	0.566	0.103	0.138	0.122
Freq_Net	-0.01	0.033	-0.02	0.039	0.015	0.294	0.05	0.043	0.023	0.033	0.026	0.071	0.087	0.31	0.027	0.01	0.102	0.087	-0.022	0.251	0.112
Freq_Fitness	0.075	0.042	0.309	0.13	0.097	-0.04	-0.052	0.081	0.14	0.166	0.335	0.221	0.124	0.068	0.132	0.138	0.173	0.163	0.267	0.085	0.086
Freq_Target	0.012	0.086	0.007	-0.067	-0.037	0.273	0.358	0.062	-0.006	-0.078	0.018	0.021	0.322	0.253	-0.017	0.009	0.014	-0.021	-0.01	0.175	0.227
Variety_Outdoor	0.277	0.766	0.215	0.106	0.24	0.204	0.189	0.707	0.254	0.091	0.195	0.198	0.188	0.224	0.635	0.307	0.247	0.167	0.223	0.225	0.236
Variety_Aquatic	0.806	0.295	0.188	0.21	0.447	0.242	0.206	0.28	0.763	0.208	0.246	0.426	0.218	0.257	0.345	0.716	0.406	0.28	0.272	0.266	0.212
Variety_Ind	0.359	0.226	0.398	0.313	0.738	0.306	0.245	0.214	0.358	0.293	0.299	0.666	0.295	0.306	0.251	0.37	0.568	0.333	0.337	0.349	0.296
Variety_Dance	0.268	0.109	0.207	0.788	0.305	0.199	0.166	0.08	0.249	0.755	0.162	0.338	0.226	0.237	0.106	0.224	0.247	0.655	0.113	0.173	0.145
Variety_Target	0.122	0.174	0.16	0.055	0.218	0.458	0.592	0.116	0.066	0.024	0.122	0.119	0.515	0.386	0.09	0.079	0.158	0.067	0.129	0.336	0.403
Variety_Net	0.234	0.2	0.287	0.104	0.268	0.716	0.474	0.149	0.179	0.05	0.248	0.217	0.449	0.615	0.101	0.149	0.268	0.09	0.24	0.505	0.378
Variety_Fitness	0.221	0.231	0.627	0.153	0.327	0.307	0.147	0.229	0.164	0.111	0.599	0.313	0.289	0.354	0.222	0.16	0.333	0.119	0.493	0.29	0.254
Identify_Outdoor	0.313	0.809	0.278	0.159	0.215	0.047	0.163	0.766	0.302	0.181	0.287	0.248	0.214	0.12	0.654	0.327	0.252	0.188	0.259	0.144	0.21
Identify_Target	0.268	0.18	0.238	0.106	0.327	0.506	0.81	0.148	0.216	0.105	0.259	0.295	0.69	0.437	0.097	0.192	0.26	0.108	0.24	0.359	0.494
Identify_Net	0.315	0.137	0.325	0.183	0.331	0.78	0.583	0.145	0.258	0.177	0.308	0.342	0.58	0.724	0.104	0.224	0.359	0.181	0.315	0.587	0.442
Identify_Fitness	0.262	0.286	0.763	0.224	0.364	0.202	0.159	0.315	0.246	0.224	0.739	0.425	0.339	0.309	0.284	0.205	0.381	0.219	0.584	0.25	0.267
Identify_Aquatic	**0.868**	0.265	0.218	0.296	0.429	0.238	0.246	0.263	0.842	0.328	0.261	0.451	0.269	0.278	0.329	0.769	0.409	0.376	0.253	0.268	0.23
Identify_Ind	0.429	0.204	0.389	0.357	0.817	0.359	0.333	0.204	0.424	0.356	0.377	0.772	0.392	0.348	0.283	0.431	0.696	0.405	0.433	0.439	0.392
Identify_Dance	0.3	0.094	0.198	**0.889**	0.341	0.208	0.187	0.091	0.304	**0.856**	0.177	0.369	0.251	0.257	0.123	0.272	0.303	0.736	0.124	0.2	0.181
Present_Outdoor	0.316	**0.864**	0.288	0.115	0.237	0.081	0.177	0.799	0.332	0.139	0.273	0.232	0.198	0.132	0.671	0.361	0.232	0.18	0.263	0.135	0.192
Present_Aquatic	**0.898**	0.311	0.217	0.274	0.444	0.209	0.208	0.281	**0.877**	0.31	0.224	0.424	0.207	0.23	0.343	0.796	0.363	0.338	0.228	0.2	0.153
Present_Ind	0.465	0.247	0.403	0.356	0.825	0.353	0.312	0.247	0.512	0.37	0.401	0.797	0.386	0.376	0.295	0.504	0.67	0.424	0.412	0.374	0.316
Present_Dance	0.293	0.124	0.213	**0.881**	0.333	0.172	0.191	0.116	0.296	**0.851**	0.154	0.337	0.244	0.219	0.131	0.262	0.241	0.706	0.093	0.127	0.124
Present_Target	0.243	0.224	0.238	0.098	0.309	0.588	0.808	0.184	0.204	0.135	0.266	0.271	0.749	0.546	0.149	0.214	0.268	0.156	0.285	0.418	0.55
Present_Fitness	0.278	0.273	0.813	0.203	0.399	0.267	0.23	0.287	0.266	0.251	0.763	0.462	0.386	0.349	0.279	0.249	0.358	0.234	0.597	0.231	0.251
Present_Net	0.319	0.127	0.304	0.2	0.354	0.828	0.632	0.115	0.289	0.222	0.293	0.341	0.595	0.744	0.111	0.279	0.332	0.226	0.307	0.561	0.434
Assess_Outdoor	0.341	**0.925**	0.35	0.163	0.286	0.23	0.247	0.846	0.335	0.156	0.294	0.288	0.279	0.288	0.71	0.369	0.295	0.205	0.282	0.224	0.28
Assess_Aquatic	**0.932**	0.387	0.306	0.354	0.509	0.318	0.313	0.331	**0.887**	0.357	0.28	0.472	0.302	0.323	0.392	0.813	0.443	0.394	0.292	0.285	0.254

	SkillAbility_Outdoor	SkillAbility_Fitness	SkillAbility_Aquatic	SkillAbility_Ind	SkillAbility_Dance	SkillAbility_Net	SkillAbilty_Target	Freq_Outdoor	Freq_Aquatic	Freq_Ind	Freq_Dance	Freq_Net	Freq_Fitness	Freq_Target	Variety_Outdoor	Variety_Aquatic	Variety_Ind	Variety_Dance	Variety_Target	Variety_Net
Assess_Dance	-0.007	0.073	0.16	0.282	0.76	0.069	-0.002	-0.048	0.067	0.166	0.635	0.03	0.133	-0.073	0.095	0.237	0.319	0.794	0.027	0.085
Assess_Fitness	0.125	0.534	0.17	0.335	0.228	0.229	0.092	0.133	0.038	0.146	0.145	-0.065	0.261	0.004	0.182	0.224	0.366	0.182	0.165	0.288
Assess_Ind	0.094	0.252	0.323	0.655	0.368	0.215	0.12	0.02	0.186	0.429	0.196	0.047	0.124	-0.025	0.176	0.444	0.753	0.329	0.186	0.254
Assess_Net	0.054	0.222	0.128	0.223	0.196	0.617	0.409	0.022	0.073	0.105	0.084	0.306	-0.018	0.258	0.16	0.233	0.27	0.198	0.446	0.669
Assess_Target	0.085	0.143	0.082	0.273	0.135	0.43	0.495	0.012	0.017	0.069	0.023	0.074	0.015	0.354	0.149	0.215	0.258	0.169	0.601	0.492
Feedback_Aquatic	0.276	0.257	0.764	0.363	0.374	0.289	0.226	0.124	0.404	0.261	0.182	-0.01	0.075	0.012	0.277	0.806	0.359	0.268	0.122	0.234
Feedback_Outdoor	0.687	0.304	0.315	0.293	0.204	0.194	0.227	0.486	0.158	0.248	-0.005	0.033	0.042	0.086	0.766	0.295	0.226	0.109	0.174	0.2
Feedback_Fitness	0.149	0.576	0.18	0.342	0.242	0.256	0.102	0.15	0.003	0.16	0.161	-0.02	0.309	0.007	0.215	0.188	0.398	0.207	0.16	0.287
Feedback_Dance	-0.018	0.1	0.126	0.263	0.744	0.087	0	-0.075	0.024	0.143	0.648	0.039	0.13	-0.067	0.106	0.21	0.313	0.788	0.055	0.104
Feedback_Ind	0.14	0.317	0.335	0.676	0.347	0.259	0.194	0.03	0.145	0.399	0.189	0.015	0.097	-0.037	0.24	0.447	0.738	0.305	0.218	0.268
Feedback_Net	0.097	0.233	0.161	0.249	0.224	0.645	0.449	0.033	0.071	0.126	0.095	0.294	-0.04	0.273	0.204	0.242	0.306	0.199	0.458	0.716
Feedback_Target	0.107	0.121	0.107	0.251	0.163	0.471	0.548	-0.02	-0.005	0.04	0.041	0.05	-0.052	0.358	0.189	0.206	0.245	0.166	0.592	0.474
Diff_Outdoor	0.643	0.314	0.282	0.268	0.207	0.145	0.15	0.498	0.179	0.257	0.042	0.043	0.081	0.062	0.707	0.28	0.214	0.08	0.116	0.149
Diff_Aquatic	0.27	0.238	0.732	0.395	0.408	0.248	0.171	0.155	0.415	0.31	0.18	0.023	0.14	-0.006	0.254	0.763	0.358	0.249	0.066	0.179
Diff_Dance	0.002	0.109	0.158	0.266	0.735	0.058	0.017	-0.048	0.055	0.167	0.613	0.033	0.166	-0.078	0.091	0.208	0.293	0.755	0.024	0.05
Diff_Fitness	0.155	0.577	0.251	0.313	0.274	0.228	0.079	0.138	0.09	0.179	0.114	0.026	0.335	0.018	0.195	0.246	0.299	0.162	0.122	0.248
Diff_Ind	0.132	0.3	0.307	0.599	0.41	0.209	0.082	0.09	0.196	0.44	0.223	0.071	0.221	0.021	0.198	0.426	0.666	0.338	0.119	0.217
Diff_Target	0.141	0.274	0.125	0.252	0.26	0.429	0.428	0.028	0.001	0.074	0.113	0.087	0.124	0.322	0.188	0.218	0.295	0.226	0.515	0.449
Diff_Net	0.126	0.291	0.178	0.234	0.299	0.558	0.37	0.089	0.083	0.159	0.149	0.31	0.068	0.253	0.224	0.257	0.306	0.237	0.386	0.615
SSN_Outdoor	0.563	0.309	0.369	0.338	0.245	0.1	0.129	0.417	0.226	0.287	0.083	0.027	0.132	-0.017	0.635	0.345	0.251	0.106	0.09	0.101
SSN_Aquatic	0.308	0.229	0.69	0.393	0.381	0.222	0.203	0.116	0.422	0.329	0.166	0.01	0.138	0.009	0.307	0.716	0.37	0.224	0.079	0.149
SSN_Ind	0.166	0.321	0.314	0.531	0.329	0.228	0.104	0.069	0.194	0.371	0.156	0.102	0.173	0.014	0.247	0.406	0.568	0.247	0.158	0.268
SSN_Dance	0.063	0.138	0.229	0.348	0.67	0.105	0.081	0.006	0.127	0.195	0.566	0.087	0.163	-0.021	0.167	0.28	0.333	0.655	0.067	0.09
SSN_Fitness	0.174	0.493	0.252	0.385	0.211	0.229	0.117	0.152	0.113	0.215	0.103	-0.022	0.267	-0.01	0.223	0.272	0.337	0.113	0.129	0.24
SSN_Net	0.107	0.296	0.2	0.318	0.215	0.421	0.295	0.052	0.112	0.226	0.138	0.251	0.085	0.175	0.225	0.266	0.349	0.173	0.336	0.505
SSN_Target	0.146	0.281	0.158	0.305	0.2	0.334	0.337	0.034	0.067	0.167	0.122	0.112	0.086	0.227	0.236	0.212	0.296	0.145	0.403	0.378

	Variety_Fitness	Identify_Outdoor	Identify_Target	Identify_Net	Identify_Fitness	Identify_Aquatic	Identify_Ind	Identify_Dance	Present_Outdoor	Present_Aquatic	Present_Ind	Present_Dance	Present_Target	Present_Fitness	Present_Net	Assess_Outdoor	Assess_Aquatic	Assess_Dance	Assess_Fitness	Assess_Ind	Assess_Net	Assess_Target
Assess_Dance	0.12	0.176	0.085	0.188	0.213	0.335	0.369	**0.894**	0.136	0.309	0.383	**0.891**	0.127	0.211	0.237	0.175	0.392	1	0.216	0.435	0.258	0.225
Assess_Fitness	0.613	0.221	0.223	0.297	0.717	0.236	0.366	0.152	0.259	0.24	0.393	0.177	0.256	0.782	0.319	0.345	0.334	0.216	1	0.465	0.401	0.377
Assess_Ind	0.309	0.176	0.298	0.328	0.356	0.442	0.823	0.339	0.188	0.447	0.842	0.323	0.29	0.391	0.371	0.262	0.532	0.435	0.465	1	0.416	0.407
Assess_Net	0.313	0.051	0.516	0.783	0.225	0.228	0.344	0.182	0.077	0.202	0.328	0.163	0.596	0.288	0.832	0.232	0.321	0.258	0.401	0.416	1	0.74
Assess_Target	0.221	0.174	0.785	0.606	0.212	0.245	0.348	0.159	0.188	0.212	0.329	0.158	0.812	0.279	0.636	0.29	0.326	0.225	0.377	0.407	0.74	1
Feedback_Aquatic	0.221	0.313	0.268	0.315	0.262	**0.868**	0.429	0.3	0.316	**0.898**	0.465	0.293	0.243	0.278	0.319	0.341	**0.932**	0.355	0.318	0.459	0.308	0.32
Feedback_Outdoor	0.231	0.809	0.18	0.137	0.286	0.265	0.204	0.094	**0.864**	0.311	0.247	0.124	0.224	0.273	0.127	**0.925**	0.387	0.185	0.308	0.253	0.188	0.266
Feedback_Fitness	0.627	0.278	0.238	0.325	0.763	0.218	0.389	0.198	0.288	0.217	0.403	0.213	0.238	0.813	0.304	0.35	0.306	0.229	**0.908**	0.471	0.338	0.332
Feedback_Dance	0.153	0.159	0.106	0.183	0.224	0.296	0.357	**0.889**	0.115	0.274	0.356	**0.881**	0.098	0.203	0.2	0.163	0.354	**0.954**	0.234	0.405	0.231	0.204
Feedback_Ind	0.327	0.215	0.327	0.331	0.364	0.429	0.817	0.341	0.237	0.444	0.825	0.333	0.309	0.399	0.354	0.286	0.509	0.404	0.477	**0.886**	0.401	0.405
Feedback_Net	0.307	0.047	0.506	0.78	0.202	0.238	0.359	0.208	0.081	0.209	0.353	0.172	0.588	0.267	0.828	0.23	0.318	0.245	0.364	0.414	**0.926**	0.695
Feedback_Target	0.147	0.163	0.81	0.583	0.159	0.246	0.333	0.187	0.177	0.208	0.312	0.191	0.808	0.23	0.632	0.247	0.313	0.198	0.302	0.36	0.655	**0.864**
Diff_Outdoor	0.229	0.766	0.148	0.145	0.315	0.263	0.204	0.091	0.799	0.281	0.247	0.116	0.184	0.287	0.115	0.846	0.331	0.171	0.292	0.218	0.174	0.245
Diff_Aquatic	0.164	0.302	0.216	0.258	0.246	0.842	0.424	0.304	0.332	**0.877**	0.512	0.296	0.204	0.266	0.289	0.335	**0.887**	0.359	0.24	0.453	0.233	0.261
Diff_Dance	0.111	0.181	0.105	0.177	0.224	0.328	0.356	**0.856**	0.139	0.31	0.37	**0.851**	0.135	0.251	0.222	0.156	0.357	**0.909**	0.207	0.364	0.194	0.195
Diff_Fitness	0.599	0.287	0.259	0.308	0.739	0.261	0.377	0.177	0.273	0.224	0.401	0.154	0.266	0.763	0.293	0.294	0.28	0.178	0.761	0.385	0.305	0.311
Diff_Ind	0.313	0.248	0.295	0.342	0.425	0.451	0.772	0.369	0.232	0.424	0.797	0.337	0.271	0.462	0.341	0.288	0.472	0.417	0.47	0.834	0.36	0.379
Diff_Target	0.289	0.214	0.69	0.58	0.339	0.269	0.392	0.251	0.198	0.207	0.386	0.244	0.749	0.386	0.595	0.279	0.302	0.277	0.402	0.398	0.619	0.804
Diff_Net	0.354	0.12	0.437	0.724	0.309	0.278	0.348	0.257	0.132	0.23	0.376	0.219	0.546	0.349	0.744	0.288	0.323	0.308	0.388	0.378	0.829	0.632
SSN_Outdoor	0.222	0.654	0.097	0.104	0.284	0.329	0.283	0.123	0.671	0.343	0.295	0.131	0.149	0.279	0.111	0.71	0.392	0.176	0.287	0.269	0.163	0.21
SSN_Aquatic	0.16	0.327	0.192	0.224	0.205	0.769	0.431	0.272	0.361	0.796	0.504	0.262	0.214	0.249	0.279	0.369	0.813	0.321	0.224	0.446	0.257	0.257
SSN_Ind	0.333	0.252	0.26	0.359	0.381	0.409	0.696	0.303	0.232	0.363	0.67	0.241	0.268	0.358	0.332	0.295	0.443	0.347	0.398	0.757	0.405	0.396
SSN_Dance	0.119	0.188	0.108	0.181	0.219	0.376	0.405	0.736	0.18	0.338	0.424	0.706	0.156	0.234	0.226	0.205	0.394	0.775	0.225	0.422	0.251	0.23
SSN_Fitness	0.493	0.259	0.24	0.315	0.584	0.253	0.433	0.124	0.263	0.228	0.412	0.093	0.285	0.597	0.307	0.282	0.292	0.104	0.593	0.427	0.345	0.326
SSN_Net	0.29	0.144	0.359	0.587	0.25	0.268	0.439	0.2	0.135	0.2	0.374	0.127	0.418	0.231	0.561	0.224	0.285	0.235	0.306	0.408	0.664	0.515
SSN_Target	0.254	0.21	0.494	0.442	0.267	0.23	0.392	0.181	0.192	0.153	0.316	0.124	0.55	0.251	0.434	0.28	0.254	0.191	0.332	0.358	0.538	0.632

	Feedback_Aquatic	Feedback_Outdoor	Feedback_Fitness	Feedback_Dance	Feedback_Ind	Feedback_Net	Feedback_Target	Diff_Outdoor	Diff_Aquatic	Diff_Dance	Diff_Fitness	Diff_Ind	Diff_Target	Diff_Net	SSN_Outdoor	SSN_Aquatic	SSN_Ind	SSN_Dance	SSN_Fitness	SSN_Net	SSN_Target
Assess_Dance	0.355	0.185	0.229	**0.954**	0.404	0.245	0.198	0.171	0.359	**0.909**	0.178	0.417	0.277	0.308	0.176	0.321	0.347	0.775	0.104	0.235	0.191
Assess_Fitness	0.318	0.308	**0.908**	0.234	0.477	0.364	0.302	0.292	0.24	0.207	0.761	0.47	0.402	0.388	0.287	0.224	0.398	0.225	0.593	0.306	0.332
Assess_Ind	0.459	0.253	0.471	0.405	**0.886**	0.414	0.36	0.218	0.453	0.364	0.385	0.834	0.398	0.378	0.269	0.446	0.757	0.422	0.427	0.408	0.358
Assess_Net	0.308	0.188	0.338	0.231	0.401	**0.926**	0.655	0.174	0.233	0.194	0.305	0.36	0.619	0.829	0.163	0.257	0.405	0.251	0.345	0.664	0.538
Assess_Target	0.32	0.266	0.332	0.204	0.405	0.695	**0.864**	0.245	0.261	0.195	0.311	0.379	0.804	0.632	0.21	0.257	0.396	0.23	0.326	0.515	0.632
Feedback_Aquatic	1	0.41	0.309	0.352	0.518	0.336	0.354	0.348	**0.883**	0.35	0.286	0.445	0.338	0.334	0.41	0.798	0.409	0.383	0.279	0.296	0.28
Feedback_Outdoor	0.41	1	0.367	0.212	0.323	0.217	0.279	**0.888**	0.386	0.198	0.28	0.309	0.287	0.258	0.742	0.389	0.291	0.22	0.258	0.201	0.25
Feedback_Fitness	0.309	0.367	1	0.272	0.504	0.359	0.317	0.336	0.239	0.26	0.777	0.512	0.411	0.379	0.318	0.211	0.441	0.247	0.613	0.29	0.314
Feedback_Dance	0.352	0.212	0.272	1	0.422	0.258	0.24	0.187	0.319	**0.891**	0.184	0.408	0.281	0.29	0.162	0.27	0.323	0.744	0.086	0.221	0.194
Feedback_Ind	0.518	0.323	0.504	0.422	1	0.462	0.437	0.275	0.45	0.385	0.409	0.816	0.47	0.439	0.314	0.438	0.702	0.435	0.416	0.446	0.407
Feedback_Net	0.336	0.217	0.359	0.258	0.462	1	0.713	0.197	0.261	0.213	0.305	0.391	0.641	**0.855**	0.176	0.275	0.416	0.271	0.317	0.682	0.548
Feedback_Target	0.354	0.279	0.317	0.24	0.437	0.713	1	0.203	0.248	0.172	0.271	0.325	0.783	0.574	0.149	0.242	0.309	0.21	0.264	0.449	0.606
Diff_Outdoor	0.348	**0.888**	0.336	0.187	0.275	0.197	0.203	1	0.427	0.261	0.392	0.389	0.35	0.348	**0.85**	0.441	0.378	0.302	0.383	0.352	0.369
Diff_Aquatic	**0.883**	0.386	0.239	0.319	0.45	0.261	0.248	0.427	1	0.439	0.326	0.535	0.333	0.352	0.492	**0.872**	0.481	0.478	0.334	0.348	0.312
Diff_Dance	0.35	0.198	0.26	**0.891**	0.385	0.213	0.172	0.261	0.439	1	0.265	0.484	0.359	0.366	0.275	0.373	0.381	0.827	0.183	0.299	0.267
Diff_Fitness	0.286	0.28	0.777	0.184	0.409	0.305	0.271	0.392	0.326	0.265	1	0.562	0.518	0.473	0.408	0.298	0.503	0.291	0.736	0.415	0.443
Diff_Ind	0.445	0.309	0.512	0.408	0.816	0.391	0.325	0.389	0.535	0.484	0.562	1	0.55	0.54	0.417	0.482	0.793	0.504	0.522	0.512	0.463
Diff_Target	0.338	0.287	0.411	0.281	0.47	0.641	0.783	0.35	0.333	0.359	0.518	0.55	1	0.783	0.299	0.321	0.49	0.392	0.425	0.614	0.745
Diff_Net	0.334	0.258	0.379	0.29	0.439	**0.855**	0.574	0.348	0.352	0.366	0.473	0.54	0.783	1	0.316	0.345	0.508	0.412	0.414	0.775	0.645
SSN_Outdoor	0.41	0.742	0.318	0.162	0.314	0.176	0.149	**0.85**	0.492	0.275	0.408	0.417	0.299	0.316	1	0.607	0.539	0.425	0.549	0.48	0.473
SSN_Aquatic	0.798	0.389	0.211	0.27	0.438	0.275	0.242	0.441	**0.872**	0.373	0.298	0.482	0.321	0.345	0.607	1	0.594	0.553	0.465	0.47	0.432
SSN_Ind	0.409	0.291	0.441	0.323	0.702	0.416	0.309	0.378	0.481	0.381	0.503	0.793	0.49	0.508	0.539	0.594	1	0.554	0.701	0.699	0.633
SSN_Dance	0.383	0.22	0.247	0.744	0.435	0.271	0.21	0.302	0.478	0.827	0.291	0.504	0.392	0.412	0.425	0.553	0.554	1	0.377	0.508	0.481
SSN_Fitness	0.279	0.258	0.613	0.086	0.416	0.317	0.264	0.383	0.334	0.183	0.736	0.522	0.425	0.414	0.549	0.465	0.701	0.377	1	0.642	0.654
SSN_Net	0.296	0.201	0.29	0.221	0.446	0.682	0.449	0.352	0.348	0.299	0.415	0.512	0.614	0.775	0.48	0.47	0.699	0.508	0.642	1	**0.858**
SSN_Target	0.28	0.25	0.314	0.194	0.407	0.548	0.606	0.369	0.312	0.267	0.443	0.463	0.745	0.645	0.473	0.432	0.633	0.481	0.654	**0.858**	1

Appendix G

Total Variance Explained-Unrotated Ten-Factor Model

Total Variance Explained- Unrotated Ten-Factor Model

	Initial Eigenvalues			Extraction Sums of Squared Loadings		
Factor	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	20.305	32.231	32.231	20.125	31.944	31.944
2	7.081	11.239	43.470	6.877	10.916	42.861
3	6.210	9.858	53.328	6.016	9.550	52.410
4	4.245	6.738	60.066	4.054	6.434	58.845
5	4.011	6.366	66.432	3.804	6.037	64.882
6	2.541	4.033	70.465	2.326	3.692	68.574
7	2.367	3.757	74.222	2.149	3.411	71.986
8	1.915	3.039	77.261	1.628	2.584	74.570
9	1.473	2.339	79.599	1.212	1.923	76.493
10	1.353	2.147	81.746	.966	1.533	78.026
11	.874	1.387	83.133			
12	.800	1.269	84.402			
13	.777	1.234	85.636			
14	.663	1.053	86.689			
15	.594	.943	87.632			
16	.552	.876	88.508			
17	.529	.839	89.347			
18	.511	.812	90.159			
19	.447	.710	90.868			
20	.405	.643	91.512			
21	.389	.618	92.129			
22	.329	.522	92.651			
23	.318	.504	93.155			
24	.295	.468	93.624			
25	.280	.445	94.068			
26	.257	.408	94.476			
27	.243	.386	94.862			
28	.221	.351	95.212			
29	.207	.329	95.541			
30	.188	.298	95.839			

Appendix H

Total Variance Explained for Unrotated Eight-Factor Model

Total Variance Explained-Unrotated Eight-Factor PAF

	Initial Eigenvalues			Extraction Sums of Squared Loadings		
Factor	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	20.305	32.231	32.231	20.100	31.905	31.905
2	7.081	11.239	43.470	6.832	10.845	42.750
3	6.210	9.858	53.328	6.003	9.529	52.278
4	4.245	6.738	60.066	4.033	6.402	58.680
5	4.011	6.366	66.432	3.771	5.986	64.666
6	2.541	4.033	70.465	2.305	3.658	68.324
7	2.367	3.757	74.222	2.109	3.348	71.672
8	1.915	3.039	77.261	1.564	2.483	74.155
9	1.473	2.339	79.599			
10	1.353	2.147	81.746			
11	.874	1.387	83.133			
12	.800	1.269	84.402			
13	.777	1.234	85.636			
14	.663	1.053	86.689			
15	.594	.943	87.632			
16	.552	.876	88.508			
17	.529	.839	89.347			
18	.511	.812	90.159			
19	.447	.710	90.868			
20	.405	.643	91.512			
21	.389	.618	92.129			
22	.329	.522	92.651			
23	.318	.504	93.155			
24	.295	.468	93.624			
25	.280	.445	94.068			
26	.257	.408	94.476			
27	.243	.386	94.862			
28	.221	.351	95.212			
29	.207	.329	95.541			
30	.188	.298	95.839			

Appendix I

Oblim Rotation of Six-Factor Varimax PAF

Pattern Matrix for Oblim Rotation: Six-Factor PAF [a]

	Factor					
	1	2	3	4	5	6
Identify_Ind	.880					
Variety_Ind	.840					
Tran_Assess_Ind	-.811					
Tran_FB_Ind	-.802					
Present_Ind	.789					
SkillAbility_Ind	.731					
Trans_Diff_Ind	-.668					
SSN_Ind	.655					
Freq_Ind	.520					
Tran_FB_Target		.833				
Tran_Assess_Target		.819				
Tran_FB_Net		.790				
Present_Target		-.790				
Tran_Assess_Net		.786				
Trans_Present_Net		.785				
Tran_Identify_Target		.775				
Tran_Identify_Net		.753				
Trans_Diff_Target		.728				
Tran_Diff_Net		.667				
Variety_Target		-.635				
SkillAbilty_Target		-.625				
Tran_Variety_Net		.596				
SkilAbility_Net		-.560				
Freq_Target		-.533				
Tran_SSN_Target		.527				
Tran_SSN_Net		.493				
Freq_Net						
Present_Outdoor			.919			
Feedback_Outdoor			.908			
Assess_Outdoor			.901			
Diff_Outdoor			.864			

Pattern Matrix for Oblim Rotation: Six-Factor PAF [a]

	Factor					
	1	2	3	4	5	6
Identify_Outdoor			.859			
Variety_Outdoor			.845			
SkillAbility_Outdoor			.776			
SSN_Outdoor			.696			
Freq_Outdooor			.588			
Present_Aquatic				.962		
Identify_Aquatic				.915		
Assess_Aquatic				.886		
SkillAbility_Aquatic				.883		
Feedback_Aquatic				.880		
Variety_Aquatic				.870		
Diff_Aquatic				.834		
SSN_Aquatic				.717		
Freq_Aquatic				.516		
Present_Dance					.961	
Identify_Dance					.953	
Assess_Dance					.941	
Feedback_Dace					.937	
Diff_Dance					.909	
Variety_Dance					.853	
SkillAbility_Dance					.765	
Freq_Dance					.747	
SSN_Dance					.723	
Tran_Identify_Fit						.858
Tran_Present_Fitness						.842
Tran_Assess_Fitness						.813
Trans_Diff_Fitness						.804
Tran_FB_Fitness						.787
Tran_Variety_Fitness						.701
SkillAbility_Fitness						-.673
Tran_SSN_Fit						.624
Trans_Freq_Fit						.465

Extraction Method: Principal Axis Factoring.
Rotation Method: Oblimin with Kaiser Normalization.[a]

a. Rotation converged in 13 iterations.

Appendix K

Descriptive Statistics for Self-Efficacy Results

	N	Mean	Std. Deviation
(0-6 scale)			
SkillAbility_Outdoor	182	3.52	1.333
SkillAbility_Fitness	182	4.79	.929
SkillAbility_Aquatic	182	3.90	1.307
SkillAbility_Ind	182	4.21	1.142
SkillAbility_Dance	182	3.68	1.279
SkilAbility_Net	182	4.82	.776
SkillAbilty_Target	182	4.62	.883
(0-5 scale)			
Freq_Outdooor	182	2.45	.994
Freq_Aquatic	182	2.48	.985
Freq_Ind	182	2.55	1.327
Freq_Dance	182	2.30	1.108
Freq_Net	182	2.84	1.043
Freq_Target	182	2.75	.952
Freq_Fitness	182	4.17	1.002
Variety_Outdoor	182	2.87	1.154
Variety_Aquatic	182	3.03	1.214
Variety_Ind	182	3.47	.990
Variety_Dance	182	2.98	1.189
Variety_Target	182	3.74	.876
Variety_Fitness	182	3.97	.834
Variety_Net	182	3.95	.782
(0-10 scale)			
Identify_Outdoor	182	6.47	2.522
Identify_Aquatic	182	7.05	2.950
Identify_Ind	182	8.06	2.265
Identify_Dance	182	6.68	2.943
Identify_Fitness	182	9.29	1.657

Identify_Net	182	9.29	1.512
Identify_Target	182	8.97	1.799
Present_Outdoor	182	6.67	2.643
Present_Aquatic	182	6.99	3.119
Present_Ind	182	8.01	2.407
Present_Dance	182	6.82	3.081
Present_Target	182	9.00	1.866
Present_Net	182	9.24	1.700
Present_Fitness	182	9.34	1.673
Assess_Outdoor	182	6.92	2.808
Assess_Aquatic	182	7.41	3.122
Assess_Dance	182	7.16	3.101
Assess_Fitness	182	9.51	1.611
Assess_Ind	182	8.54	2.278
Assess_Net	182	9.47	1.713
Assess_Target	182	9.32	1.765
Feedback_Outdoor	182	6.92	2.802
Feedback_Aquatic	182	7.37	3.061
Feedback_Dance	182	7.09	3.108
Feedback_Fitness	182	9.42	1.696
Feedback_Ind	182	8.48	2.224
Feedback_Net	182	9.37	1.712
Feedback_Target	182	9.14	1.839
Diff_Outdoor	182	6.96	2.884
Diff_Aquatic	182	7.26	3.172
Diff_Dance	182	7.07	3.191
Diff_Fitness	182	9.45	1.713
Diff_Ind	182	8.46	2.347
Diff_Target	182	9.15	1.926
Diff_Net	182	9.30	1.848
SSN_Fitness	182	8.83	2.192
SSN_Net	182	8.83	2.105

SSN_Target	182	8.75	2.149
SSN_Outdoor	182	6.76	3.062
SSN_Aquatic	182	6.87	3.272
SSN_Ind	182	8.20	2.611
SSN_Dance	182	7.09	3.048
Valid N (listwise)	182		