PERSONALIZED SYSTEM OF INSTRUCTION AND STUDENT PERFORMANCE IN HIGH SCHOOL WEIGHT TRAINING COURSES

by

Charles Allen

Liberty University

A Dissertation Presented in Partial Fulfillment

Of the Requirements for the Degree

Doctor of Education

Liberty University

July, 2015

PERSONALIZED SYSTEM OF INSTRUCTION AND STUDENT PERFORMANCE IN HIGH SCHOOL WEIGHT TRAINING COURSES

by Charles P. Allen

A Dissertation Presented in Partial Fulfillment

Of the Requirements for the Degree

Doctor of Education

Liberty University, Lynchburg, VA

July, 2015

APPROVED BY:

John Pantana, Ed.D, Committee Chair

Steven McDonald, Ed.D, Committee Member

Yesenia M. Lopez, Ed.D, Committee Member

Scott B. Watson, PhD, Associate Dean of Advanced Programs

ABSTRACT

School reform has applied pressure on the United States public school systems to improve student achievement. As a result of this pressure, educators are seeking instructional models that research supports improves student achievement. The purpose of this causal comparative study was to test the Social Cognitive Theory by comparing the achievement of all-male high school weight training students who had been taught using the *personalized system of instruction* (PSI) instructional model to students who had not been taught using the PSI instructional model on the state mandated Fitnessgram assessments, after controlling for prior Fitnessgram achievement within a large, urban high school in northeast Georgia. Archival Fitnessgram pretest and posttest data was collected on a total of 206 students, of which 103 having been taught by teachers using the PSI instructional model and 103 having been taught without the PSI instructional model. The data collected was then analyzed by ANCOVA to determine the possible effect of instructional model on student achievement on the Fitnessgram PACER, ninety degree push-up, and curl-up assessments, after controlling for prior Fitnessgram achievement as measured by the Fitnessgram pretest scores. The data revealed no statistically significant difference in student achievement between the groups on any of the Fitnessgram assessments and each of the null hypotheses were not rejected. Suggestions for further research are included.

Keywords: personalized system of instruction, physical education, student achievement, Fitnessgram

Dedication

This study is dedicated to my Lord and Savior, Jesus Christ. Without his blessings this project would not have been possible. All of the glory goes to Him, who gave his life for all so that we might be redeemed to the Father. I would also like to dedicate this study to my mother, Deborah, who taught me how to love learning, how to work, and how to dream. I would also like to dedicate this study to my wife, Lauren, whose constant encouragement, support, sacrifice, and faith inspires me daily to be the best man I can be and to run the race the Lord has set before me with a smile on my face and a hope for the future.

TABLE OF CONTENTS

List of Tables

List of Figures

List of Abbreviations

Adequate Yearly Progress (AYP)

Analysis of Covariance (ANCOVA)

Common Core State Standards (CCSS)

Direct Instruction (DI)

Elementary and Secondary Education Act (ESEA)

Georgia Performance Standards (GPS)

Georgia Technical Committee (TAC)

Institute of Education Sciences (IES)

Institutional Review Board (IRB)

Multivariate Analysis of Variance (MANOVA)

National Assessment of Educational Progress (NAEP)

National Commission on Excellence in Education (NCEE)

No Child Left Behind Act (NCLB)

Personalized System of Instruction (PSI)

Non-Personalized System of Instruction (non-PSI)

Race to the Top (RT3)

Standard Error Measurement (SEM)

Student Growth Percentile Measures (SGP)

Student Health and Physical Education (SHAPE)

Student Performance Goals (SPG)

Teacher Keys Effectiveness System (TKES)

Trends in International Mathematics and Science Study (TIMSS)

United States Department of Education (U.S. DOE)

What Works Clearinghouse (WWC)

CHAPTER 1: INTRODUCTION

The purpose of this chapter is to establish a framework for this proposed causal comparative study. This chapter is organized as follows: (a) background, (b) problem statement, (c) purpose statement, (d) the significance of the study, (e) the research questions, (f) the research hypotheses, (g) the identification of the variables, (h) the definition of key terms.

Background

Public education in the United States during the last half of the 20th Century has been filled with public criticism and reform. According to experts, (Dufour, Dufour, & Eaker, 2008; Grady, 2009, Waite, 2000) many of these criticisms have been spurred by world events such as the 1957 launching of Sputnik, the rise of the economic and industrial power of Japan, the results of the 1999, 2003, and 2007 *Trends in International Mathematics and Science Study*, as well as national reports on the poor standards of education in the United States. An example of one of these reports is entitled *A Nation at Risk* (National Commission on Excellence in Education, 1983). This report has been cited by many (National Council of Teachers of Mathematics, 1989; National Science Foundation, 1988; Pixler, 2009; Sizemore, 2010; Ward, 2009) as being one of the main reasons for the current reform movement in education. In the report, the NCEE (1983) proposed that the U.S. educational system was "being eroded by a rising tide of mediocrity" (p.5) within its schools. The report advised the raising of educational standards. Following this report, many national organizations began the process of developing national standards for their subject areas. This process of developing standards and raising the bar in education was further emphasized with the passing into law of the *No Child Left Behind* (NCLB) *Act*.

With its focus on standardized testing and accountability for results, NCLB is one of the most controversial educational reform acts that have been passed. Signed into law by President George W. Bush in January, 2002, NCLB was the latest iteration of the *Elementary and Secondary Education Act* (ESEA), the primary federal law of public education in the United States. Though most widely known for its emphasis on high stakes, standardized testing, NCLB had four main focuses. These focuses include:

- Accountability for results,
- An emphasis on doing what works based on scientific research,
- Expanded parental options, and
- Expanded local control (U.S. Department of Education, 2005, para. 2).

The first focus, accountability for results, was built on the foundation of standardized testing and holding schools accountable for student results through the use of Adequate Yearly Progress (AYP) markers. The second focus with its emphasis on best practices helped to develop national organizations such as the What Works Clearinghouse (WWC) whose primary job is to identify studies "that provide credible and reliable evidence of the effectiveness of a given practice, program, or policy" (Institute of Education Sciences, n.d., para. 2). The third focus, expanded parental options, gave parents whose students were in low-performing schools the ability to transfer to another, higher performing school (U.S. DOE, 2005). The final focus of NCLB allowed schools and school districts greater flexibility in exchange for higher accountability for results. Though each of these contributed to raising the bar for education in the U.S. (Wilson, 2012), it was with NCLBs focus on standardized testing and higher accountability for schools that helped to further *A Nation at Risk*'s charge to raise standards in public education by driving the development of grade level and subject area state and national standards (Rand, 2008).

While subjects such as mathematics began developing standards for learning as early as 1989 (NCTM, 1989, p.1), many other subject areas lagged behind in their development until the NCLB catalyst (Rand, 2008). Physical education was one of the last subject areas to develop standards, finally publishing national standards in 1995 (National Association of Sport and Physical Education, 1995). In Georgia, the Student Health and Physical Education (SHAPE) Act was passed in 2009. The SHAPE act set forth two standards beginning in the 2011-2012 school year. These standards include the following:

- Students in grades 1 through 12 will enroll in a physical education class and receive an annual physical fitness assessment, and
- The results will be collected so that it may aid future policy decisions (Georgia Department of Education, 2009).

Proscribed benefits include establishing baseline data, tracking and monitoring trends in health related fitness over time, establishing the possibility for linkages for other indicators, and enabling the development of data driven strategies to combat childhood obesity (Georgia Department of Education, 2010).

In June of the following year, the Fitnessgram was chosen by the Georgia Department of Education (GaDOE, 2010) as the annual physical fitness assessment for students. The Fitnessgram is a "comprehensive health-related physical fitness and activity assessment and computerized reporting system" (GaDOE, 2010). Components of health-related fitness that are measured by the Fitnessgram in the state of Georgia include aerobic capacity as measured by either the 1-mile run or the PACER, muscular strength and endurance as measured by the 90 push-ups and the curl-up, flexibility as measured by the back-saver sit and reach, and body composition as measured by BMI, skinfold measurements, or bioelectric impedance analyzers (Kinetics, 2013).

Along with the SHAPE act, Georgia's commitment to the Race to the Top (RT3) federal initiative has continued to raise the bar for physical education and more specifically for physical education teachers. RT3 is a $4.35 billion competitive grant program provided by the American Recovery and Reinvestment Act of 2009 that encourages and rewards "States that are creating the conditions for education innovation and reform" (GCPS, 2014a). One of the methods Georgia is using to race to the top is through the development of the teacher evaluation system, or Teacher Keys Effectiveness System (TKES). The TKES is an extensive system geared towards measuring teacher effectiveness using a variety of formative and summative assessments on the Teacher Assessment on Performance Standards (TAPS), as well as data collected on student growth and academic achievement (GCPS, 2014a). TAPS include a series of ten research-based performance standards that are assessed through teacher observations; while student growth and academic achievement is measured using either student growth percentile (SGP) measures or student performance goal (SPG) measures. SGPs are used in situations where state mandated assessments are used to measure student achievement. These include fourth through eighth grade criterion referenced competency tests (CRCTs) as well as high school end-of-course-tests (EOCTs). Student performance goals (SPGs) are used to "measure growth in student achievement for teachers of non-state-tested subjects (GCPS, 2014b, pg.2). Physical education courses fall into this category of assessment. The Fitnessgram assessment is currently the SPG used in most physical education courses and a multiple choice SPG is under development to supplement the Fitnessgram assessment.

Along with this increased focus on teacher accountability, the recession of the last several years has greatly impacted the landscape of public education, further intensifying the pressures on school districts and teachers. According to the Center on Budget and Policy Priorities (2013), over the course of the last six years at least thirty-four states have cut funding for public education. The state of Georgia is not immune to these budgetary problems. Since 2008, Georgia's per-student funding has dropped 14.8% (The Century Foundation, 2014). These cuts have resulted in decreased funding per student, teacher layoffs, increased classroom sizes, less spending on teacher and student materials including textbooks, cuts in funding for elective courses, as well as an increase in safety concerns for students (Kelly, 2014; The Century Foundation, 2014). These budgetary problems have been particularly devastating to physical education departments and courses, specifically with regard to classroom sizes. According to the Georgia Department of Education (2012) the maximum class size for a physical education course without a paraprofessional has ballooned to forty students while a physical education course with a paraprofessional is now fifty-four students. The funding class size for most other subject areas in grades 9-12 is twenty-three (Georgia Department of Education, 2012).

Given the tremendous amount of pressure on schools imposed by NCLB to increase student achievement and meet adequate yearly progress (AYP), "districts must select curriculum, instructional, and assessment methods that help students demonstrate increased knowledge and skills on state assessments" (Pixler, 2009, p. 4). This pressure is further intensified and focused on teachers with the recent implementation of the

Teacher Keyes Effectiveness System, as well as budgetary cuts to public education that has decreased per-student funding and increased class size. One method that is being investigated in raising student achievement in physical education classes is the use of the personalized system of instruction (PSI) instructional model.

The personalized system of instruction (PSI) model is a student-centered instructional model that enables students to work at their own pace to master skills and progress through prescribed learning tasks with the teacher acting as a facilitator, tutor, and motivator, rather than as the primary source of knowledge (Metzler, 2005). According to Metzler (2005), PSI is designed to encourage independent learning for students while also allowing the teacher greater freedom to interact with students who need extra support (p. 219). The model was originally designed by Keller (1968) for use in an introductory psychology course of over 300 students. Keller decided the traditional classroom model of lecturing would not be effective for a course of that size and he set out to develop an instructional method that would "provide an individual learning program for all students" (Metzler, 2005, p. 217). As a colleague of the famous behavior psychologist B.F. Skinner and with a background in applied behavior analysis and experimental behavioral psychology, Keller believed that the whole classroom environment, not just the teacher, impacted student learning and if this were true then it should be possible to design a learning environment that could promote student learning with or without direct instruction from the teacher (Metzler, 2005). While the PSI model was originally designed for psychology courses of large sizes, certain findings suggest PSI could generate positive effects in other educational fields (Cregger, 1994; Cregger & Metzler, 1992; Eppler & Ironsmith, 2004; Hannon, Holt, & Hatten, 2008; Hansen, Brothen, & Wanbach, 2002; Leech, 2011; Lowry & Thornburg, 1988; Pritchard & Colquitt, 2006; Pritchard, Penix, Colquitt, & McCollum, 2012).

An example of PSI improving student learning and achievement in a high school physical education setting is a study conducted by Hannon, Holt, and Hatton in 2008, entitled *Personalized system of instruction model: Teaching health-related fitness content in high school physical education.* In the study, 26 students enrolled in a high school physical education weight training course were taught over three weeks a unit on post-rehabilitation using the PSI instructional model. Data was collected using audio-visual equipment, student and teacher observations, as well as a student survey using a Likert scale. Observation data was coded independently by two trained graduate students as well as the researchers and the inter-rater reliability for frequency and duration coding was found to be in acceptable range (93-97%). Researchers found that 93.4 % of students met or exceeded performance criteria. Researchers concluded that based on the confirmation criteria developed by Cregger and Metzler (1992) a PSI model could be successfully implemented in a physical education weight training course with a high degree of success for students.

Another example of PSI improving student learning and achievement was a study conducted by Pritchard, Penix, Colquitt, and McCollum in 2012, entitled *Effects of a weight training personalized system of instruction on fitness levels and knowledge*. In the study, the researchers used Fitnessgram assessment and a fifty question knowledge test as a pre and post- test assessment to measure the effectiveness of PSI in a fifteen-week beginning university physical education weight training course. The Fitnessgram assessment included the progressive aerobic cardiovascular endurance run (PACER) test, back-saver sit and reach test, trunk lift test, push test, and percentage body fat test. The fifty question knowledge test (McGee & Farrow, 1987) was designed to assess overall weight training knowledge. Participants included 17 male and 5 female university students with an age range from 18 years to 48 years, ($M = 20.77$, $SD = 6.24$). A paired-samples t test with a Bonferoni correction was used to compare pre- and post-test scores. Researchers found a statistically significant difference in the pre- and post-test scores for the curl-up test, push-up test, percentage body fat test, and knowledge test. There was no statistically significant difference found between the pre- and post-test scores on the PACER, back-saver sit and reach, or the trunk lift tests. Researchers concluded that the PSI model was effective in raising achievement.

Recent comparisons of PSI to other instructional methods are limited (Metzler, 2005). Taveggia (1976) reviewed 14 comparative studies of PSI to conventional teaching methods in higher education from several disciplines and found the PSI courses to be superior (p.1032). Kulik (1976) reviewed 31 studies comparing PSI to conventional teaching methods and found that of the studies, 25 of them produced favorable results for PSI. More recent meta-analysis studies by Kulik et al (1990) compared exam scores of PSI with Bloom's Learning for Mastery model and found that of the 67 studies reviewed, 62 of them reported higher final exam scores for students who received the PSI instructional model, with an effect size of .48 (p. 292).

While findings such as these suggest that the PSI model has the potential to improve student achievement in physical education courses, many teachers continue to resist implementing the model. Critics point to the decline in the use of the model since the 1970's as evidence of its inapplicability to today's educational landscape (Leech, 2011), while others (Buskist et al, 1991; Sherman, 1992) point to the incredible amount of initial development time required for PSI courses, difficulty in adapting the self-paced/mastery model to academic calendars, resistance by educators to transition to a student-centered approach from the traditional teacher-centered one, and the tendency of administrators not to value the PSI model as reasons not to adopt the model. Another criticism of the model is that much of the research done on PSI is dated at least fifteen years old (Leech 2011) and that the landscape of education has so vastly changed that much of the research is not applicable to today's educational system. Other critics point to the lack of research on the model as it pertains to the physical education setting as even strong proponents of the model such as Metzler (2005) admit that PSI research is limited in that arena. Still others point to the lack of research in a high school physical education setting as reasons to be hesitant to implement the model.

In reviewing the above studies as well as the criticisms, it becomes apparent that there exists a gap in the research on the personalized system of instruction instructional model. Specifically, more research is necessary to explore the possible impact of PSI on student achievement in a high school physical education setting as measured by the state-mandated Fitnessgram assessment. Furthermore, the foundation of the PSI model lies in social cognitive theory, which posits that learning is a product of psychological and environmental factors. Consequently, there exists a gap in the research as it pertains to social cognitive theory, specifically as it pertains to environmental factors that help shape student learning and achievement in high school physical education classes.

Problem Statement

The pressures on schools and school districts to meet Adequate Yearly Progress (AYP) makes it imperative that districts "select curriculum, instructional, and assessment methods that help students demonstrate increased knowledge and skills on state assessments" (Pixler, 2009, p. 4). This pressure is further intensified and focused on teachers with the recent implementation of the Teacher Keys Effectiveness System, as well as budgetary cuts to public education that has decreased per-student funding and increased class sizes. The personalized system of instruction (PSI) is an instructional model that has been examined by many studies with varying degrees of success and scope (Cregger, 1994; Cregger & Metzler, 1992; Eppler & Ironsmith, 2004; Hannon, Holt, & Hatten, 2008; Hansen, Brothen, & Wanbach, 2002; Leech, 2011; Lowry & Thornburg, 1988; Pritchard & Colquitt, 2006; Pritchard, Penix, Colquitt, & McCollum, 2012). The problem is that while many schools and educators are considering implementing a PSI instructional model in high school physical education classes in hopes of improving student achievement on the state mandated Fitnessgram assessment, there is very little current research on the effectiveness of the PSI model in high school physical education classes in raising student achievement. Furthermore, with the PSI model affecting the classroom environment, there also exists a gap in the research of social cognitive theory and how it pertains to environmental factors influencing student achievement in high school physical education classes.

Purpose Statement

The purpose of this causal comparative study is to test the social cognitive theory that relates the instructional model received by a student to student achievement, controlling for prior student achievement for high school physical education students. The independent variable for this study is the type of physical education instructional model a student receives and will be generally defined as a personalized system of instruction (PSI) instructional model or a non-personalized system of instruction (NPSI) instructional model. The dependent variable will generally be defined as student scores on the state mandated Fitnessgram assessments, and the control variable, prior student achievement on the Fitnessgram assessments will be statistically controlled in this study. By exploring the possible impact of the personalized system of instruction (PSI) on high school student achievement, this study will contribute to the body of research on PSI in high school physical education courses in large, urban school systems. Furthermore, the theory I will use is social cognitive theory. The theory is largely attributed to Albert Banduras and it is used to study human learning and behavior (Boston University, 2013). The theory proposes that human learning and behavior is a product of psychological and environmental factors (para. 1). As applied to my study, this theory holds that I would expect the instructional model used in a classroom to influence student achievement of high school students on the Fitnessgram assessments because the instructional model is a form of environmental change that the theory posits would “automatically lead to changes in the person(‘s),” (Boston University, 2013, para. 4) learning and behavior.

Significance of the Study

The importance of this study is multi-faceted. With the pressures imposed on schools by *No Child Left Behind* (NCLB) to show adequate yearly progress (AYP), it is imperative that school officials “select curriculum, instructional, and assessment methods that help students demonstrate increased knowledge and skills on state assessments” (Pixler, 2009, p. 4). This pressure is further intensified and focused on individual teachers with the recent changes in the teacher evaluation system due to the Race to the Top (RT3) federal initiative. These changes in the teacher evaluation system gauge teacher effectiveness by several factors, including student growth and achievement on student performance goal measures (GCPS, 2014a). Along with this pressure, recent decreases in funding for public education has resulted in a decrease in per-student funding and increased class sizes (Kelly, 2014; The Century Foundation, 2014). Current research on the possible impact of PSI on student achievement in high school physical education courses is limited (Hannon, Holt, & Hatton, 2008; Metzler, 2005). This study will seek to fill this gap in research on the impact of PSI on student achievement in high school physical education classes. Furthermore, this study will provide information to administrators and physical education course decision makers on the effectiveness of PSI in improving student achievement in large class-sized physical education classes. This will aid decision makers in organizing high school physical education courses. Additionally, for high school physical education teachers who are currently using the PSI model, this study will provide insight into their physical education course structure and aid them improving their own PSI courses. Along with this, the study will contribute to the growing literature on social cognitive theory and how it may apply in high school physical education settings.

Research Questions

This causal comparative study has been designed to answer several questions. These questions include the following:

RQ1: Is there a statistically significant difference between the progressive aerobic cardiovascular endurance run (PACER) scores on the state mandated Fitnessgram assessment for high school all-male physical education weight training students who have been taught using the personalized system of instruction instructional model and those who have not been taught using the personalized system of instruction instructional model, after adjusting for prior achievement on the PACER test?

RQ2: Is there a statistically significant difference between the 90 degree push-up scores on the state mandated Fitnessgram assessment for high school all-male physical education weight training students who have been taught using the personalized system of instruction instructional model and those who have not been taught using the personalized system of instruction instructional model, after adjusting for prior achievement on the 90 degree push-up test?

RQ3: Is there a statistically significant difference between curl-up scores on the state mandated Fitnessgram assessment for high school all-male physical education weight training students who have been taught using the personalized system of instruction instructional model and those who have not been taught using the personalized system of instruction instructional model, after adjusting for prior achievement on the curl-up assessment?

Research Hypotheses

Below is a description of the null hypotheses associated with each of the above research questions. Each hypothesis is presented as a null hypothesis. The null hypotheses include the following:

$H_0$1: There will be no statistically significant difference in achievement for high school all-male physical education weight training students who have been taught using the personalized system of instruction instructional model and those students who have not been taught using the personalized system of instruction instructional model as measured by the Fitnessgram progressive aerobic cardiovascular endurance run (PACER) scores, after adjusting for differences in students' prior PACER achievement.

$H_0$2: There will be no statistically significant difference in achievement for high school all-male physical education weight training students who have been taught using the personalized system of instruction instructional model and those students who have not been taught using the personalized system of instruction instructional model as measured by the Fitnessgram 90 degree push-up scores, after adjusting for differences in students' prior 90 degree push-up achievement.

$H_0$3: There will be no statistically significant difference in achievement for high school all-male physical education weight training students who have been taught using the personalized system of instruction instructional model and those students who have not been taught using the personalized system of instruction instructional model as measured by the Fitnessgram curl-up scores, after adjusting for differences in students' prior curl-up achievement.

Identification of Variables

There are several variables involved in this study. Below is a description of the independent, dependent, and control variables.

Independent variables. The operational definition of the categorical independent variable associated with this study is the instructional model students received. Possible values for the independent variable include: personalized system of instruction (PSI) or non- personalized system of instruction (NPSI). A student's instructional model will be defined as PSI if the student is enrolled in a physical education course where the PSI model is used as the primary instructional model. A student's instructional model will be defined as NPSI if the student is enrolled in a physical education course where the personalized system of instruction is not used as the as the primary instructional model. This definition of the independent variable is consistent with many studies of this nature (Cregger, 1994; Hannon, Holt, & Hatten, 2008; Metzler, 1986). A teacher survey will be used as the primary method of identifying and confirming the instructional model used. An example of the teacher survey can be found in Appendix C. Along with the survey, teacher lesson plans, pacing guides, and student manuals will also be used to confirm the instructional model used in the course.

Dependent variable. The operational definition of the dependent variable is the student's post-test scores on the Fitnessgram assessment. These assessments are designed to measure health-related fitness that includes aerobic capacity, body composition, muscular strength, endurance, and flexibility (Kinetics, 2014). Assessments include scores on the progressive aerobic cardiovascular endurance run (PACER), the 90 degree push-up assessment, and the curl-up assessment. The PACER score is a discrete ratio variable with scores ranging from 0 to 300 (Kinetics, 2014). The 90 degree push-up assessment score is a discrete ratio variable with scores ranging from 0 to 99 (Kinetics, 2014). The curl-up assessment score is a discrete ratio variable with scores ranging from 0 to 75 (Kinetics, 2014). The definitions of the above dependent variables are consistent with several studies of this nature (Floate, 2011; Roberts, 2009; Wilson, 2012; Woodward, 2009).

Control variable. The operational definition of the control variable for this study is the student pretest scores on the Fitnessgram assessment. As mentioned above, these assessments are designed to measure health related fitness (Kinetics, 2014). Assessments include scores on the progressive aerobic cardiovascular endurance run (PACER), the 90 degree push-up assessment, and the curl-up assessment, The PACER score is a discrete ratio variable with scores ranging from 0 to 300 (Kinetics, 2014). The 90 degree push-up assessment score is a discrete ratio variable with scores ranging from 0 to 99 (Kinetics, 2014). The curl-up assessment score is a discrete ration variable with scores ranging from 0 to 75 (Kinetics, 2014). Precedence for using the Fitnessgram pretest scores as a control variable for a study of this nature can be found in the research performed by Zachary Wilson (2012).

Definitions

Fitnessgram. A fitness assessment that measures aerobic capacity, muscular strength and endurance, flexibility, and body composition using a battery of tests (Kinetics, 2014), and is a state-mandated physical education assessment for Georgia (Georgia Department of Education, 2014).

Personalized System of Instruction (PSI). A research-based, student-centered instructional model developed by Fred Keller in 1968 (Metzler, 2005; Pritchard, Penix, Colquitt, & McCollum, 2012) that has the following characteristics:

- Go-at-your-own pace,
- Unit perfection requirement,
- Use of lectures and demonstrations as vehicles of motivation,
- Related stress upon the written word in teacher-student communication, and
- Use of proctors to allow repeated testing, immediate scoring, tutoring, and a marked enhancement of personal-social aspect of the educational process (Pritchard et al., 2012).

Race to the Top (RT3). A "competitive grant program designed to encourage and reward States that are creating the conditions for education innovation and reform" (GCPS, 2014a). In Georgia, a byproduct of this program has been a redesigning of the teacher evaluation system into the Teacher Keys Effectiveness System.

Student Performance Goals (SPGs). A metric used under the new teacher evaluation system to measure teacher effectiveness. SPGs assess student growth for teachers of non-tested courses, such as physical education, in Georgia. (GCPS, 2014a).

Teacher Keys Effectiveness System (TKES). An extensive system geared towards measuring teacher effectiveness using a variety of formative and summative assessments on the Teacher Assessment on Performance Standards (TAPS), as well as data collected on student growth and academic achievement (GCPS, 2014a).

Teacher Assessment on Performance Standards (TAPS). A series of ten research-based performance standards that are assessed through teacher observations. These standards include professional knowledge, instructional planning, instructional strategies, differentiated instruction, assessment strategies, assessment uses, positive learning environment, academically challenging environment, professionalism, and communications (GCPS, 2014a).

CHAPTER 2: LITERATURE REVIEW

Introduction

The following review of literature will explore the pertinent literature related to this study. This review of literature is organized in the following manner: (a) theoretical framework, (b) review of pertinent literature, and (c) summary. The theoretical framework explores social cognitive theory and how the theory relates to this study. The review of pertinent literature is organized into several major themes that lead us to the natural development of the current study. These themes include the following:

1. Review of the current state of high school physical education in the United States, standards, reform, and obstacles
2. Gender-grouping in physical education
3. Model-based instruction in physical education
4. The PSI instructional model
5. PSI in physical education

Reviewing the current state of high school physical education explores the standards and reform efforts along with the challenges that physical education instructors are facing. Gender grouping in physical education investigates one of the methods physical education instructors are using to overcome their current challenges. In this section, a brief history of gender-grouping in education can be found along with research on the impact of gender-grouping in physical education. Model-based instruction in physical education reviews the history of instructional practices in physical education along with the eight instructional models for physical education set forth by Metzler (2005a). The next theme, the PSI instructional model, explores the key features of PSI along with a brief history on the rise and fall of the model as well as research into the model's effectiveness. Finally, PSI in physical education reviews all the literature on PSI in a physical education setting. The review of literature concludes with a summary of key points, an identification of gaps in the research, and an explanation of how the study seeks to fill these gaps.

Theoretical Framework

The goal of this research is to explore the possible impact of the *personalized system of instruction* (PSI) instructional model on student performance on the state mandated Fitnessgram assessment for all-male physical education weight training students. At the core of this research is a change in student environment from a teacher-centered, instructional model to a student-centered instructional model. A theory that suggests such a change would affect student learning is Social Cognitive Theory.

Social Cognitive Theory (SCT) is a theory of human behavior that is largely attributed to Albert Bandura (1977). The theory is an expansion of Montgomery's Social Learning Theory, which was developed in the late 1800s, that theorized human behavior is a product of only cognitive factors. SCT on the other hand posits that human behavior and knowledge acquisition is a product of the interactions between current behavior with environmental and psychological factors (Denler, Wolter, & Benson, 2014). Furthermore, SCT theorizes that human learning often occurs in a social environment and through observing others modeling behaviors. Through these observations, individuals form expectations about consequences for specific behaviors (PSU, n.d.).

The instructional practices, goals, domain priorities, behavior models, and overall classroom environment offered in a *personalized system of instruction* (PSI) classroom will differ dramatically from those offered in the traditional teacher-centered classroom. It is the central hypothesis of this study that student behavior in a PSI classroom will adapt to this new environment and will result in positive, statistically significant results on the state mandated *Fitnessgram* assessment when compared with students in the traditional classroom, after controlling for prior achievement.

Review of the Literature

The Current State of Physical Education in the United States

Physical education, similar to public education as a whole in the U.S., over the last several decades can be summed up in two words: testing and reform. With the passing into law of the *No Child Left Behind* (NCLB) legislation in January of 2002, high stakes testing, or tests whose "results are used to make important decisions that immediately affect students, teachers, administrators, communities, schools, and districts (Au, 2009, p. 44) has become a standard in education. U.S. Senator Paul D. Wellstone (2002), an educator for twenty years prior to taking office states the following with regard to high stakes testing:

> When used correctly, standardized tests are critical for diagnosing inequality and for identifying where we need improvement. They enable us to measure achievement across groups of students so that we can help ensure that states and districts are held accountable for improving the achievement of all students regardless of race, income, gender, limited English proficiency and disability… Using a single standardized test as the sole determinant for graduation, promotion, tracking and ability grouping is not fair and has not fostered greater equality or opportunity for students. (para 9)

With regard to reform movements over the past fifty years, "a case could be made the nation has engaged in a continuous, unabated, even frenzied effort to improve its schools" (Dufour, Dufour, & Eaker, 2008, p. 45).

Standards for Physical Education

One product of these reform movements has been the development of subject area standards and a raising of accountability for schools and teachers in helping students reach these standards, or at the very least, to perform well on the standardized tests that measure these standards. Physical education has not been exempt from these reforms. In 1986, the National Association of Sport and Physical Education (NASPE) began to develop a definition of a physically educated person. The result of this five year project was to define a physically educated person as one who:

1. Has learned skills necessary to perform a variety of physical skills.
2. Does participate regularly in physical activity.
3. Is physically fit.
4. Knows the implications of and the benefits from involvement in physical activity.
5. Values physical activity and its contributions to a healthful lifestyle (NASPE, 1992).

Following this definition, content standards for physical education began to be developed. In 1995, NASPE published the book *Moving into the Future: National Standards for Physical Education* that set forth seven contents standards for physical education. A second edition of the book released in 2004 revised and reduced the standards to six. Following several revisions, the Society of Health and Physical Educators America (SHAPE America) set forth the following five standards that develop a framework for a quality physical education program:

- Standard 1: The physically literate individual demonstrates competency in a variety of motor skills and movement patterns.

- Standard 2: The physically literate individual applies knowledge of concepts, principles, strategies and tactics related to movement and performance.
- Standard 3: The physically literate individual demonstrates the knowledge and skills to achieve and maintain a health-enhancing level of physical activity and fitness.
- Standard 4: The physically literate individual exhibits responsible personal and social behavior that respects self and others.
- Standard 5: The physically literate individual recognizes the value of physical activity for health, enjoyment, challenge, self-expression and/or social interaction (SHAPE America, 2013)

To supplement these standards, the American Alliance for Health, Physical Education, Recreation and Dance (AAHPERD) Curriculum Framework Task Force developed grade-level outcomes that demonstrate competency in the above standards. These outcome standards are organized by grade level and school level including elementary grades K-5, middle grades 6-8, and high school 9-12. For high school students, outcomes for the standards are organized into two levels with the first level indicating the minimum knowledge and skills to be learned for college/career readiness, and the second level allowing students to build on these minimum knowledge and skills (2013).

Following the initial development of national standards, the state of Georgia began developing standards that would align with those set forth by the NASPE. Along with these standards, in 2009 the state passed the Georgia Student Health and Physical Education (SHAPE) Act. The SHAPE Act was a collaborative effort between the Governor's Office, the Georgia Department of Education, Children's Healthcare of Atlanta, The Georgia Department of Community Health, Division of Public Health, the

Atlanta Falcons Foundation, and the Department of Education Fitness Advisory Committee. Georgia's vision for the Act was "to be a nation-wide model for the use of a standardized fitness assessment in schools, and to develop data-driven strategies to address childhood obesity" (GADOE, 2009, para 4). The Act required that beginning in the 2011-2012 school year, each local school district conduct an annual fitness assessment for all students enrolled in a physical education course taught by a certified physical education instructor starting in first grade (GADOE, 2009). In June of 2010, the GADOE chose the *Fitnessgram* as the physical fitness assessment.

Testing, Reform, and Raising the Accountability for Educators.

Along with developing standards and establishing a state mandated assessment for physical education, Georgia's Department of Education was also committing itself to the Race to the Top (RT3) federal initiative. The RT3 initiative is a competitive grant program provided by the American Recovery and Reinvestment Act of 2009 that encourages and rewards states for their innovation in the following educational reform areas:

- Recruiting, preparing, rewarding, and retaining effective teachers and principals, especially where they are needed most;
- Adopting standards and assessments that prepare students to succeed in college and the workplace and to compete in the global economy;
- Building data systems that measure student growth and success, and inform teachers and principals about how they can improve instruction;
- Turning around our lowest-achieving schools (GADOE, 2014).

Georgia was awarded $400 million from the federal government to implement its RT3 plan. The GADOE has partnered with 26 school systems around the state to implement its plan, using each districts Title 1 formula for the dispersion of a portion of the funds (2014). According to the GADOE (2014), Georgia's RT3 plan includes initiatives that address data systems to support instruction, great teachers and leaders, improving early learning outcomes, innovation fund, science, technology, engineering, and mathematics (STEM), standards and assessments, and turning around lowest achieving schools. One of the methods Georgia is using to address these initiatives is through the development of its teacher and leader evaluation system.

The Teacher Keys Effectiveness System (TKES) is the latest iteration of Georgia's teacher evaluation system. The system is focused on measuring a teacher's effectiveness in two distinct methods. Each of the two methods is weighted as 50% of a teacher's overall effectiveness measure (TEM). The first method consists of a series of formal and informal teacher observations, called the Teacher Assessment on Performance Standards (TAPS). The purpose of TAPS is to measure a teacher's performance on ten criteria that include professional knowledge, instructional planning, instructional strategies, differentiated instruction, assessment strategies, assessment uses, positive learning environment, academically challenging environment, professionalism, and communication. The second method involves measuring a student's growth and academic achievement and is broken into two distinct categories for teachers: teachers of tested subjects and teachers of non-tested subjects. Teachers of tested subjects are measured for effectiveness using the Student Growth Percentile (SGP). The SGP is a growth model that uses test data collected over multiple years as pretest scores and the end of the year test as its post test score. The end of the year test varies by grade level and by school level. For instance, a fifth grade student will take the fifth grade Criterion-Referenced-Competency-Test (CRCT) at the end of the school year while a high school student may take an end-of-course-test (EOCT) at the end of a semester. Teachers of non-tested subjects are measured for effectiveness using the Student Performance Goals (SPG). The SPG is a tool used to quantify student growth of non-tested subjects which involves a pretest given at the beginning of the school year and a post-test which is given at the end of the school year. Physical education courses fall into this category of assessment with the *Fitnessgram* being used as both a pretest and post-test measurement for students and teachers.

Along with this increase in teacher accountability for student results, as previously mentioned student growth and achievement is weighted as 50% of a teacher's effective measure (TEM), the recession of the past several years has been devastating to many school districts. Districts across the country have experienced a decrease in per student spending, teacher layoffs, increased class sizes, as well as cuts in funding for electives courses (Kelly, 2014; The Century Foundation, 2014). Physical education has been particularly devastated by these cuts with regard to class size. In Georgia, the maximum class size for a physical education course without a paraprofessional has risen to forty students while a course with a paraprofessional is now set at fifty-four students (GADOE, 2012). Funding for class sizes for most other subject areas in grades 9-12 is 23 (GADOE, 2012). Given the current state of physical education, with high accountability and large classes sizes, many educators are looking for methods that can be implemented that research supports can have a positive impact on student learning. One of these methods, gender-grouping, is discussed below.

Gender-Grouping in Physical Education Courses

One of the ways physical education teachers are attempting to overcome this higher accountability for student performance on the *Fitnessgram,* while at the same time having significantly larger class sizes, is through gender-grouping students in certain physical education courses. The use of single-gender classes in the U.S. is not a new practice. Prior to 1900, education was largely a single-sex, all-male endeavor (Bradley, 2009) but with the changing norms and societal views, gradually a "coeducational model became not only evident, but necessary" (p.2). Single-gender public schools and classrooms continued to persist in certain areas of the country until the passing of Title IX of the Educational Amendment of 1972, which made the practice illegal. Recent legislation changes provided under NCLB has allowed the use of single-gender classrooms once again. These regulations published on October 25, 2006 allow public schools to offer single-gender classrooms if the school (1) provides a rationale for offering a single-gender class in that subject, (2) provides a coeducational class in the same subject at a geographically accessible location, and (3) conducts a review every two years to determine if the single-gender class is still necessary (National Association for Single Sex Public Education, 2013).

As one might expect, gender grouping in a physical educational setting can be somewhat controversial. Proponents of coeducational physical education courses claim that these classes provide equal opportunity for participation and interaction for both sexes (Koca, 2009), as well as opportunities for all students to improve cooperation and empathy skills. Furthermore, these proponents contend that differences between the sexes, such as possible differences in motor skills or muscular strength and endurance, are not an issue and that coeducational courses guarantee equal opportunities for both genders (Pfister, 2005). Others argue that enrollment in a coeducational physical education course does not guarantee equality and that many other variables including instructional method used, student perceptions, and teacher interactions impact the equality of instruction (Hannon & Williams, 2008). Findings by Sadker & Sadker (1993) seem to strengthen this position of non-guaranteed equity within coeducational settings. In their three year study of over 100 schools in Connecticut, District of Columbia, Maryland, Massachusetts, and Virginia, trained observers found teachers favoring male students in many ways including: allowing males to call out answers rather than insisting on them raising their hands as the female students were required to do, valuing male comments over female comments, and through encouraging the males to solve problems on their own using critical thinking skills (1993). Other proponents for single-gender physical education courses argue that social interactions between the two sexes can negatively affect female participation and illicit unequal educational opportunities. Research completed by Olafson (2002) seems to support many of these findings. Olafson (2002), while studying female adolescent resistance to school, found that many females attempted to avoid physical education courses because of uncomfortable peer interactions and that gender segregation might be a method to improve female participation. Findings such as these suggest support for this move by many physical education teachers, including the ones participating in this study, towards single-gender physical education classes.

While the deregulation of single-gender classrooms is fairly recent and controversial at times, there are several studies of its use in a physical education setting. These studies include investigations into teacher and student perceptions, confidence levels, and preferences (Hannon & Ratliffe, 2007; Hannon & Williams, 2002; Hill, Hannon, & Knowles, 2012; Lirgg, 1993; Olafson, 2002; Sinclair, 2000), student activity and engagement levels (Gabbei, 2004; Hannon & Ratliffe, 2005; McKenzie, Prochaska, Sallis, & LaMaster, 2004; Schmitt, 2001), physical fitness and academic performance in other subject areas (Roberts, 2009; Rodenroth, 2010; Wittber, Northrup, & Cottrell, 2012; Woodward, 2009) and teacher behaviors and interactions (Hannon & Ratliffe, 2007; Lirgg, 1993; Nilges, 1998). With regard to the possible impact of single-gender physical education classes on physical fitness assessments such as the *Fitnessgram*, there is only one recent study conducted by Wilson (2012).

Wilson's 2012 study, entitled *The Effects of Single-Gender Classes on Student Attitudes and Physical Fitness Test Performance*, used the Physical Fitness Attitudinal Scale along with the *Fitnessgram* assessment to investigate the possible impact of gender-grouping on 277 sixth grade students' attitudes towards single-gender physical education classes as well as their performance in physical fitness activities. Students participated in the *Fitnessgram* pretest assessment and then were subsequently divided into an all-male group, all-female group, and a coeducational group. Students were then administered the *Fitnessgram* post-test along with the Physical Fitness Attitudinal Scale. Analyses of covariance (ANCOVA) as well as multivariate analysis of variance (MANOVA) were used to test for statistically significant differences between the groups. Findings revealed statistically significant differences in group performances in some of the *Fitnessgram* assessments. These assessments included the curl-up, push-up, and the one-mile run. Certain findings also suggested that coeducational settings for females adversely affected posttest scores on portions of the *Fitnessgram* assessment. Wilson (2014) concluded that portions of the data supported a "promising relationship between gender grouping and physical fitness assessment performance" (p.83). Findings such as these appear to support a move towards gender-grouping courses in certain situations, including physical education courses. Unfortunately, research into the use of gender-grouping within the various physical education courses such as weight training is limited.

This proposed study explores the educational landscape within single-gender weight training courses and will help to fill this gap in research.

Model-Based Instruction in Physical Education

Another method many physical education teachers are investigating to improve student achievement is through the use of model-based instruction. Model-based instruction for physical education is an emerging new method of instruction that provides physical education teachers an array of instructional models to choose from that are often times very different from the traditional, teacher-focused *sage on the stage* models that have become synonymous with physical education instruction. These research supported instructional models allow teachers to differentiate their instruction based on several important factors to teaching and learning. These factors include:

- Intended learning outcomes,
- Context and teaching environment,
- Student developmental stage and readiness,
- Student learning preferences,
- Domain priorities,
- Task structure and organizational patterns,
- Sequencing of learning tasks,
- Assessment of learning outcomes,
- Assessment of instructional practices (Metzler, 2005a, p. 17-18).

While model-based instruction appears to be a promising new methodology for physical education teachers, it has roots in four previous stages of physical education instruction.

Below is a description of these four stages followed by an explanation of model-based instruction.

Methods to Models: Five Stages of Physical Education Instruction

Over the course of the last 100 years, physical education instruction has gone through a series of five stages (Metzler, 2005a). These stages include a focus on teaching methods, teaching strategies, teaching styles, teaching skills, and most recently, instructional models (Metzler, 2005a). The first stage, with its focus on teaching method, is characteristic of early 1900's physical education training programs. During this time instructional methods tended to be direct and formal (Van Dalen & Bennett, 1971), with teachers having control of the learning environment through a series of systematic lessons that were procedure oriented (Metzler, 2005a). Lessons at this time emphasized drills and repetition whose end product for students would be a level of proficiency at a given skill or sport. Examples of activities taught during stage one includes gymnastics and some sports.

With the arrival of the 1960's, stage one began to give way to another form of physical education instruction that focused less on rigid teacher control but on teaching strategies that engaged students. During this time the student's role in the classroom became more important and a variety of strategies were used that presented more freedom for students to interact with the teacher, other students, and the content of the lesson. Examples of popular teaching strategies include task and station teaching, reflective teaching, peer teaching, team teaching, and inquiry-based teaching (Metzler, 2005a). Research completed during stage two tended to center around these instructional strategies and their effectiveness when compared to other strategies (Graham, 1981).

The next stage in physical education instruction focused on teaching styles rather than teaching strategies and was heavily influenced by Muska Mosston ‘s 1966 book *Teaching Physical Education*. In the book Mosston (1966) introduced the *Spectrum of Teaching Styles* that conceptualized styles of teaching along a continuum which progressed from teacher-centered styles that were direct and formal to student-centered styles that were considered indirect and informal. A style was placed on this continuum based upon the degree of responsibility assumed during the lesson by the teacher and the student (Doherty & Ferguson, 2010). Table 1 provides a summary of Mosston’s (1966) *Spectrum of Teaching Styles*.

Table 1
Summary of Mosston’s (1966) Spectrum of Teaching Styles

Style	Summary of the Style
Style A: Command	The teacher makes all decisions.
Style B: Practice	Teacher makes decisions and students carry out tasks assigned by the teacher.
Style C: Reciprocal	Student work in pairs. One performs a task and the other student provides feedback.
Style D: Self-Check	Students assess their own performance based on a given criteria.
Style E: Inclusion	Teachers plan the work and students monitor their own work.
Style F: Guided Discovery	Students solve movement problems proscribed by the teacher with assistance.
Style G: Divergent	Students problem solve without the support of the teacher.
Style H: Individual	Teacher determines the content of the lesson, the student plans the lesson.
Style I: Learner Initiated	Student plans their own program and the teacher advises.
Style J: Self-Teaching	Student takes complete responsibility for the learning process.

Table Note: summarized from Mosston (1966)

While Mosston’s (1966) work is now over fifty years old, its impact on physical education is still felt today. The framework for teaching physical education that the *Spectrum* established was so influential that Nixon and Locke (1973) described the work as "the most significant advance in the theory of physical education pedagogy in recent history" (p.1227).

The arrival of the 1980s saw another shift in physical education instruction. Research done at this time focused on the notion of effective teaching sparked a shift away from some of Mosston's (1966) teaching styles and towards the idea of effective teaching skills. According to Metzler (2005a), much of this research explored teacher and student behaviors that increased achievement. Correlational research findings suggested that student behavior was "more predictive of learning than teacher behavior" (Metzler, 2005a, p. 13). Consequently, physical education teachers began focusing less on teacher behavior and more on what the teacher was getting students to do in class, with an effective teaching skill being defined as any intentional decision or action that increased the possibility of a student learning in class (Metzler, 2005a).
Examples of effective teaching skills for physical education include:

- Start and stop cues
- "back to the wall"
- Instant activities,
- Use of questions
- Use of cross-group feedback (Metzler, 2005a).

While each of the above stages in physical education instruction have had distinct impacts on teaching and learning, according to Metzler (2005a) each of the stages are limited in scope and are generally used for a short time and with "a few short-term learning activities and outcomes, before giving way to another method, strategy, style, or skill" (p. 13). That being said, a new stage in physical education instruction has emerged over the last several years that builds upon the previous four stages while focusing not on methods, strategies, styles, or skills, but on instructional models.

Instructional Models of Physical Education Instruction

Model-based instruction in physical education, while having roots in the four previous stages also represents a paradigm shift in the way educators organize and plan instruction. In previous stages, much of what went on in physical education courses could be considered *activities-based instruction*. *Activities-based instruction* is the practice of allowing the activity, or content, to drive instruction. For instance, a physical education teacher might spend twenty years teaching archery the same way without any thought to the differences in the learning styles of students, student readiness to learn the content, or any other possible factor that might affect student learning simply because "I teach archery this way" (Metzler, 2005a). If this educator is a stage one methods educator he or she might spend twenty years teaching archery by organizing lessons around rigorous skill and drill techniques. If the educator was a strategist based teacher he or she might use a collection of strategies- station teaching, peer teaching, or perhaps inquiry teaching to teach the archery. Educators who use instructional-models on the other hand think very differently. These educators consider a vast array of factors, including content, before deciding on which instructional *model* to use to instruct their students. Metzler (2005a) defines an instructional *model* as

> a comprehensive and coherent plan for teaching that includes a theoretical foundation, statements of intended learning outcomes, teacher's content knowledge expertise, developmentally appropriate and sequenced learning activities, expectations for teacher and student behaviors, unique task structures,

> measures of learning outcomes, and mechanisms for assessing the faithful implementation of the model itself (p. 16).

Simply put, an instructional model is a unique blueprint that educators follow that helps them to plan, design, implement, and assess entire units of instruction (Metzler, 2005a). There are many advantages to using model-based instruction in physical education. These advantages include:

- providing an overall plan and coherent approach to teaching and learning,
- clarifying learning domain priorities and domain interactions,
- providing an instructional theme,
- allowing teachers and students to understand current and upcoming events
- furnishing a unified theoretical framework,
- is research supported,
- promotes a technical language for teachers,
- allows the relationship between instruction and learning to be verified,
- allows for more valid assessments of learning,
- encourages teacher decision making within a unified framework,
- directly promote specific standards and learning outcomes (Metzler, 2005a).

There are a total of eight instructional models that research has shown are appropriate for physical education (Metzler, 2005a). A brief description of these eight models can be found in table 2 below.

Table 2
Instructional Models for Physical Education

Instructional Model	Description
Direct Instruction	Teacher as Instructional Leader
Personalized System of Instruction	Students Progress as Fast as They Can or as Slow as They Need

Cooperative Learning	Students Learn With, By, and For Each Other
Sport Education	Learning to Become Competent, Literate, and Enthusiastic Sportspersons
Peer Teaching	"I Teach You, Then You Teach Me"
Inquiry Teaching	Learner as Problem Solver
Tactical Games	Teaching Games for Understanding
Teaching for Personal and Social Responsibility	Integration, Transfer, Empowerment, and Teacher-Student Relationships

Table Note: Adapted from Metzler (2005a)

Each of the instructional models described above is designed to have a unique foundation, teaching and learning features, and implementation needs and modifications. Within each of these is a series of features that ground each instructional model. Table 3 summarizes each of these features.

Table 3
Summary of Foundations, Teaching and Learning Features, and Implementation Needs for Model-Based Instruction

Foundations	Teaching and Learning Features	Implementation Needs
Theory and Rationale	Directness and Inclusiveness	Teacher Expertise
Assumptions about Teaching and Learning	Learning Tasks	Key Teaching Skills
Theme	Engagement Patterns	Contextual Requirements
Learning Domain Priorities	Teacher/Student Roles and Responsibilities	Contextual Modifications
Student Developmental Requirements	Verification of Instructional Processes	
Validation	Assessment of Learning	

Table Note: Adapted from Metzler (2005a)

Physical education teachers using model-based instruction in their classrooms review learning goals and domain priorities for the learning unit, compare those to the domain priorities of the eight instructional models, and then deductively decide which of the models is most appropriate for usage with their students for that particular unit (Metzler, 2005a). Following this, the teacher follows the instructional model blueprint to make instructional decisions, plan lessons, clarify teacher and student roles and responsibilities, communicate future events, clarify learning goals, assess student learning, and assess the effectiveness of implementation of the model (Metzler, 2005a).

Support for model-based instruction in physical education has grown in recent years. The implementation of model-based instruction in teacher education programs has been shown to provide a structured way to organize content in a relevant and meaningful way for student teachers, cooperating teachers, as well as K-12 students (Gurvitch, Metzler, & Lund, 2008). With regard to the instructional models themselves, empirical evidence suggests that each model can "lead to intended learning outcomes in physical education that are a part of their natural design" (Barrett, 2005; Cregger & Metzler, 1992; Dyson, 2002; Dyson, Griffin, & Hastie, 2004; Griffin & Butler, 2005; Hannon et. Al; 2008; Ward & Lee, 2005; Woods, 2007).

While the use model-based instruction in physical education has continued to grow over the last several years (Gurvitch & Metzler, 2013), with the current pressures on physical education teachers to raise student achievement while also working with large class sizes, one model in particular, called the *personalized system of instruction* (PSI) model, has recently begun to grow in popularity.

The PSI Instructional Model

The *personalized system of instruction* (PSI) instructional model is a student-centered instructional model that enables students to progress through prescribed learning tasks at their own pace to master skills set forth by the teacher (Metzler, 2005a). Cregger (1994) describes the model as "an interlocking system of instruction, consisting of sequentially progressive tasks designed as highly individualized learning activities" (p. 16). The model was primarily developed by Fred Keller in the early 1960s for use in an introductory psychology course of over 300 students at the University of Brazil. Following a presentation of B.F. Skinner's principle of Analysis of Behavior, Keller and several associates decided that "traditional teaching methods were sadly out of date" (Keller & Sherman, 1974, p. 7). Keller suggested that if education was to improve, instructional design systems would need to be developed that would update methods of providing instruction. Furthermore, Keller concluded that a methodical pattern of instruction should be used that builds upon students' previous success to reinforce progress towards a specified outcome (Cregger, 1994). Originally called the Keller Plan, Keller (1968) identified five essential features to his plan that include: (a) student self-pacing, (b) mastery learning, (c) use of lectures and demonstrations as vehicles of motivation, (d) emphasis on written word in teacher-student communication, and (e) the use of proctors for immediate student support. Over time, the system Keller developed would come to be called the *personalized system of instruction* (PSI).

The Growth and Decline of PSI

Preliminary support for PSI came very quickly. Endorsements by the National Science Foundation and the Alfred P. Sloan Foundation helped to develop and spread the plan nationally and internationally (Keller, 1974). By 1973, over 300 research articles on PSI had been published (Sherman, 1982). PSI course offerings also began to increase and to expand into a variety of disciplines. Table 4, adapted from Keller (1974) summarizes the subject areas and number of courses offered using PSI in 1972 and 1974. Notice that the total number of courses offered more than doubled from 1972 to 1974. Also note the diverse subject matter being taught using the PSI model during this time.

Table 4
Summary of the Subject Matter and Number of PSI Courses Offered in 1972 and 1974

Subject Matter	Number of Courses Offered in 1972	Number of Courses Offered in 1974
Biology	6	21
Chemistry	15	31
Engineering	21	49
English	4	11
Mathematics – Statistics	20	49
Physics	38	53
Psychology	73	157
Sociology	3	16
Other	10	23
Totals	190	410

Table Note: Adapted from Keller (1974)

By 1979, the number of PSI courses being offered in a variety of disciplines had ballooned to well over 5000 (Sherman, 1982). Furthermore, a newsletter entitled, the *PSI Newsletter*, a journal entitled, *Journal of Personalized Instruction*, and a clearinghouse for PSI course offerings located at Georgetown University, called the Center for Personalized Instruction, was established during this time frame. Along with this, PSI workshops and conferences were offered.

Research on the effectiveness of PSI in those early days was also promising. In a summary of over 100 research reports through 1973, Robin (1974) concluded that thirteen out of fifteen contrast studies favored PSI, while research by Kulik, Kulik, and Carmichael (1974) found that learning of content using PSI was adequate. Kulik et. al (1974) also compared student performance on final examinations between courses taught using PSI and courses taught using other methods and found that PSI student performance was equal and often times better than students taught using other methods. Explorations into the impact of PSI on student attitudes, creativity, self-actualization, study habits, self-image, and dropout rate were also done at this time.

Unfortunately, as is the case with many educational movements, over time support for PSI began to wane. Administrative differences and funding issues caused the Center for Personalized Instruction to become non-operational (Sherman, 1982). Along with this, confidence in the effectiveness of PSI began to decline. Four major criticisms contributed to this decline in confidence. These include the economic cost of the model, the low level of interaction between students and teachers, poor performance by students on standardized tests, and a growing debate over what level of mastery is obtained by students (Wichita, n.d.). The first criticism, the economic cost of the model was in large part due to the models dependency on specially trained and paid proctors and tutors. Along with this, the substantial amount of time that instructors needed to spend on developing the materials and training the proctors and tutors contributed to raising the cost of these courses. The second criticism, low levels of interaction between students and teachers stemmed from the role of the teacher changing from the primary source and dispenser of knowledge to manager of student learning (Gallup & Allan, 2003). Often times, the teacher would spend tremendous amounts of time developing materials and training proctors and very little time interacting with students. This lack of interaction with students did little to improve the teacher's knowledge about successful teaching practices (Silberman, 1974) and was a common complaint among PSI instructors. Another common criticism was the poor performance of PSI students on standardized tests. According to Kulick, Kulick, and Bangert-Drowns (1990), PSI students tended to do very well on instructor developed exams but poorly on standardized tests when compared to students in conventional courses. This criticism is controversial, as there are many studies of PSI that show otherwise. Regardless, this research led many to believe that PSI courses were teaching to test (Kulick, et. al, 1974). The final major complaint revolved around grading and mastery learning. Common grading practice in a PSI course was that a student worked on a unit until a mastery level was achieved and then received an A for that unit (Wichita, n.d.). Differences between levels of mastery within the content were difficult to determine and arguments over the definition of mastery ensued. Instructors often found this grading system confounding (Wichita, n.d.) and difficult to use. While the growth and decline in the popular use of the PSI can be traced back to the criticisms above along with many others, there exists a substantial amount of research on PSI and student learning and performance.

Further Research on the Effectiveness of PSI

Research over PSI is extensive (Hymel, 1987). Lowry and Thornburg (1988) cite over 1500 articles in their research while Sherman (1992) suggests over 2000 research studies have been conducted over PSI. While this number is substantial, much of the research is over twenty years old and the number of recent studies is greatly reduced (Buskist et. al, 1991; Lamal, 1984; Leech, 2011). Eyre (2007) further puts the lack of current research in perspective when she notes that less than 50 studies were performed over the sixteen year period between 1990 and 2006. With such a dated research base for PSI, more research is needed. Below is a discussion of the current research base for PSI.

Results-Based Research

One branch of research over PSI is focused on comparing the results of PSI to another instructional model or strategy, similar to this proposed study. Research over this branch of investigation into PSI is particularly dated with much of the research occurring while the Center for Personalized Instruction was still operational at Georgetown University.

There are several meta-analytical studies that summarize research over PSI's effectiveness when compared to other instructional models or strategies. As previously mentioned, work done by Robin (1974) Kulik et. al (1974) helped to lay a research foundation for PSI as a viable instructional method through their early comparison studies. Another study, conducted by Taveggia in 1976 is entitled, *Personalized instruction: A summary of comparative research, 1967-1975* also extended research over PSI. In the study Taveggia (1976) reviewed 14 studies in a number of disciplines including anthropology, chemistry, and psychology. Each of the 14 studies compared PSI with traditional teaching methods, using student scores on courses exams as variables. Taveggia (1976) concluded the PSI courses were shown to be superior. Another study completed by Kulik (1976) reviewed 31 studies that compared PSI methods to traditional teaching. Kulik (1976) found that of the 31 studies, 25 of them found significantly higher final exam scores for courses taught using PSI, while the remain six studies found no significant differences between the two instructional methods. Kulik (1976) also found higher student perceptions, retention rates, and transfer effects in those courses taught using PSI. A third meta-analysis study completed Kulik, Kulik, and Cohen (1976) compared PSI to non-PSI courses along outcomes that included course completion and withdraw rates, final exam scores, final course grades, student satisfaction, and student study time. Results of the study found that PSI outperformed the non-PSI courses in each of the above arenas.

Another meta-analysis study by Kulik, Kulik, and Cohen (1979) reviewed 72 studies that compared PSI to non-PSI instruction. The authors focused on studies that had outcome measures that used final exam scores, final course grades, student satisfaction, student study time, and course completion and withdrawal rates. A total of

75 courses were reviewed and compared in the study. Kulik et al (1979) found the final exam scores of PSI students to be 8 percentage points higher than those in the non-PSI courses, with an average effect size of .5. Information on how effect size was measured and interpreted was not given. Student retention of material was also investigated. PSI students scored 14 percentage points higher than non-PSI students. Comparing final course grades revealed similar results, with PSI students scoring nearly a full letter grade higher than non-PSI students. With regard to students satisfaction, Kulik et al (1979) writes "students rate PSI classes as more enjoyable, more demanding, and higher in overall quality and contribution to student learning than conventional classes" (p. 317).

Another study by Kulik et al (1990), reviewed 67 comparative studies of PSI to Bloom's Learning for Mastery model. Of the 67 studies reviewed, 62 of them revealed higher final exam scores for the PSI students, with an effect size of .48.

Other Avenues of PSI Research

More recent investigations into PSI tend to move away from comparative studies between PSI and other instructional methods. One avenue of recent investigation is into the applicability of the PSI model with distance learning programs. With a focus on written word, PSI appears to be well suited for use in distance learning programs and recent research suggests it could be effectively implemented (Conard, 1997; Grant & Spencer, 2003; Lui, 2003). Furthermore, with the use of computer based course management systems such as Blackboard and D2L that have the capability to reduce administrative duties, offer peer tutoring, collaborative activities, grade exams and give feedback, as well as offer opportunities for multiple attempts at mastery level achievement, PSI appears to be a viable option for this form of education (Lui, 2003).

Along with distance learning education, the expansion of web-based instruction and blended classrooms has offered interesting avenues for PSI. Research by Svenningsen and Pear (2011) on Computer-aided personalized system of instruction (CAPSI) in blended courses suggest a that PSI could be effectively implemented in these settings. In the study, Svenningsen and Pear (2011) investigated the impact of CAPSI on student course knowledge and critical thinking development. In one portion of the study, 364 University of Manitoda students enrolled in a 13-week introductory course. A total of four course sections were used with two sections using the CAPSI system and the other using traditional teaching methods. Final exam scores were used to measure student achievement. ANOVA was used to analyze data. Results revealed that CAPSI section students mean scores were 3.23 points higher than non-CAPSI section students. The effect size was measured using partial eta-squared and was classified as minimal.

While PSI appears to be well suited for distance learning and blended classrooms, other avenues of research more pertinent to this study have investigated how to fix particular problems within Keller's PSI framework. One such problem is student procrastination. Student procrastination is a natural product of the self-pacing tenant of PSI (Fox, 2004) and is not easily fixed (Eyre, 2007). Researchers have investigated several different approaches to fix this problem with varying degrees of success. These approaches include the use of behavioral contracts (Brooke & Ruthven, 1984), teaching students time-management skills (Keenan, Bono, & Hursh, 1978), using a bonus point system for completing tasks early (Eppler & Ironsmith, 2004; Reidel, Harney, LaFief, & Finch, 1976; Semb, Conyers, Spencer, & Sanchez-Sosa, 1975), using student set deadlines (Roberts & Semb, 1989; Roberts & Semb, 1990), and consequences for students not meeting deadlines (Miller, Weaver, & Semb, 1974).

Another problem that is inherent to PSI that has been recently investigated is mastery learning. If you will recall, problems with mastery learning are considered one of the reasons for the recent decline in PSI taught courses since the 1970s (Wichita, n.d.). Issues with mastery learning that have been investigated include problems with defining grading criteria (Wichita, n.d.), as well as students running out of time and not mastering the material (Eyre, 2007). Research into how to fix these problems include differentiating assessments to include PSI as a portion of the overall grade rather than the whole grade (Brothen & Wambach, 2001; Eppler & Ironsmith, 2004), limiting the number of retakes of tests of master (Eyre, 2007), and the use of a conditional pass or equivalent system (Crone-Todd, 2007; Liu, 2003).

An investigation into some of the historically systematic problems of PSI mentioned above has offered interesting methods of updating the instructional model to the 21st century educational landscape. As Sherman (1992) warned "a rigid definition (of PSI) can freeze the method into a numbing formula and limit the audience" (p. 62). Research completed by Fox (2004) investigated ways of updating the model for the 21st century while at the same time offering the model a broad since of flexibility for instructors. Table 5 summarizes these updates and revisions.

Table 5
Updated Key Features of the Personalized System of Instruction for 21st Century Education

Feature	Description
Unit Mastery	Students are required to demonstrate unit mastery before proceeding to the next unit.
Flexible Pacing	Students proceed through course content at their own pace. Teachers are encouraged to use strategies to reduce procrastination.
On-Demand Course Content	Students have access to instructional materials whenever needed; instruction material medium can vary.
Immediate Feedback	Students receive immediate feedback on assessments through

	either human or computerized means
Peer Tutoring	Peer tutoring is available to provide support, feedback, and administer assessments when necessary

Note: Adapted from Fox (2004)

In summary, a review of the above literature over PSI reveals three key points. The first point is PSI is a thoroughly researched instructional model. The second point is that research suggests PSI can be implemented in a variety of educational settings, courses, and subject areas. The final key point is that much of the research, including comparative studies between the model and other instructional models, is dated at least twenty years old, with a bulk of these studies being done while the Center for Personalized System of Instruction was still operational (Leech, 2011). One subject area whose teachers have recently shown interest in exploring the uses of the model is physical education. It is to PSI's implementation in physical education courses we will now turn to.

PSI in Physical Education Courses

As previously mentioned, physical education (PE) courses in today's educational climate have some very real challenges to overcome. Some of the more notable challenges include a lack of funding, large class sizes, and increased accountability for results. While the first two challenges are not within the domain of control for PE teachers, the last, with its focus on results, can be if teachers are willing to accept and adapt with this new PE landscape. One method PE departments are exploring to raise student achievement is the implementation of the PSI instructional model. Below is a discussion of the framework of PSI in PE courses as well as a review of current literature.

As mentioned above, the use of the PSI instructional model has expanded into a wide variety of subject areas and many believe the model to be well suited for implementation in PE courses (Colquitt, Pritchard, & McCollum, 2011; Hannon, Holt, & Hatton, 2008; Leech, 2011; Metzler, 2005a; Pritchard, Penix, Colquitt, & McCollum, 2012). While this appears to be the case, Metzler (2005a), a proponent for the use of the model-based instruction in PE courses, suggests the model to be most appropriate for middle, high school, or college PE courses that have the following criteria:

- courses with activities that can be broken into discrete skills or knowledge areas that should be learned in a definite sequence,
- courses with a strong emphasis on learning outcomes in the psychomotor domain (Metzler, 2005a, p. 239).

Along with the above criteria, Metzler (2005a) suggests that when choosing an instructional model for PE courses the instructor should be well versed in each of the eight instructional models, including their assumptions about teaching and learning. Table 6 provides a summary of these assumptions for the PSI model.

Table 6
Assumptions about Teaching and Learning for the PSI Model

Assumptions about Teaching	Assumptions about Learning
Many class management functions can be completed without the teacher	Student learning can occur with little dependence upon the teacher
The teacher's primary function is to interact with students for learning and motivation, not for class management.	Students learn at different rates.
Student engagement and learning are most effective when they remain largely independent of the teacher.	Students have differing aptitudes for learning content.
Planning decisions are driven by data collected on student learning.	All students can achieve the stated learning goals if given enough time and/or trials.
It is possible to design individualized instruction for each student.	Students will be highly motivated and responsible as independent learners.

Note: Adapted from Metzler (2005a)

Building upon the above criteria, Metzler (2005a) suggests PSI to be most appropriate for individualized sports courses, team sports courses, recreational activities courses, dance courses, personal fitness concept courses, as well as personal fitness program courses (p. 239). Weight training courses such as those investigated in this study, with an almost exclusive focus in the psychomotor domain and emphasis on sequenced teaching of individual skills, appear to be well suited for the PSI model. A review of the Academic Knowledge and Skills (AKS) for weight training courses set forth by GCPS (2014) lends support for these claims. A list of the AKS for weight training courses can be found in Table 7.

Table 7
Academic Knowledge and Skills (AKS) for High School Weight Training Courses

AKS	Reference Code
A – Weight Training for Fitness	
Demonstrate correct training methods used in weight training	PEWT_A2009-1
Identify weight loads, number of sets, and repetitions in various weight training programs	PEWT_A2009-2
Identify the types of exercises to be performed in order to enhance the development of various muscle groups	PEWT_A2009-3
Develop and plan a series of exercises in order to maximize the benefits of a weight training program	PEWT_A2009-4
Explain the importance of performing large muscle group exercises prior to small or isolated muscle group movements	PEWT_A2009-5
Describe why the altering method of push-pull or upper body-lower body exercise method is performed in order to maximize training benefits	PEWT_A2009-6
Describe the importance of determining the amount of rest needed between sets and training workout routines in order to maximized training	PEWT_A2009-7
Describe the causes and effects of over-training	PEWT_A2009-8
Identify the major muscle groups of the body	PEWT_A2009-9
Achieve and maintain a health-enhancing level of physical fitness	PEWT_A2010-1
B – Weight Training Equipment and Aids	
Identify how to properly use the two major types of weight training equipment: machines and free weight	PEWT_A2009-10
C – Program Organization and Technique	
Demonstrate proper technique in executing various lifts	PEWT_A2009-11
Describe the importance of the “warm-up” and “cool-down” phase of the training program in order to prepare the body for stress and recovery	PEWT_A2009-12

Demonstrate the proper spotting techniques for various lifts	PEWT_A2009-13
Describe the importance of charting and record-keeping in a weight training program	PEWT_A2009-14
D – Nutrition, Rest, and Ergogenic Aids	
Identify the harmful effects of anabolic steroids and dietary supplements on the body and how they affect the weight training program	PEWT_A2009-15

Note: Taken from Academic Knowledge and Skills 2014-15 High School (GCPS, 2014)

Research on PSI within Physical Education

Research on the PSI model within PE courses is limited and is often found pertaining to college-level courses (Metzler, 2005a). Support for the uses of individualized instruction in PE can be traced back to the 1970s, when PSI was gaining in popularity. Singer and Dick (1974) believed the interests and needs of individual students had largely been ignored in PE courses and those needs and interests should be considered of the upmost importance if the acquisition of motor skills was to be successful. Daryl Siedentop (1974) was the first to bring the benefits of PSI to PE when he described how college-level activity courses could use PSI. Annarino (1976), reporting results from several studies of PSI in PE settings, found that PSI results were equally or more effective than other teaching methods.

Tousignant (1983) helped to lay a foundation for PSI in a high school PE setting. In the study, Tousignant (1983) used the PSI model to teach her high school tennis class. During the study students were allowed to progress at their own pace, were given reading materials to learn about tasks, and were required to master tasks before proceeding to the next task. The results of the study concluded that PSI was effective in helping students reach the outcome criteria set forth by the teacher. Interestingly, the author also noted a need for careful planning by the teacher in order to successfully implement PSI, a sentiment later echoed by Metzler (2005a).

Several studies by Metzler furthered the research base of the PSI model within PE settings. In one study Metzler (1986) compared the effectiveness of the PSI model to a group-demonstration model. The author concluded that PSI in physical education has the potential to better attend to individual student needs, foster better use of time, allow for more student practice of motor skills, and promote greater student success (p. 7). A later study by Metzler (1988) investigated student achievement and process in a college level tennis course by comparing 8 PSI courses to 8 courses taught using conventional means. Findings of the study revealed that PSI students were more engaged, received more instructional content, were provided more practice, and had a higher rate of successful motor trials than the conventionally taught students. Further research done with the same participants found that PSI teachers gave more feedback than their non-PSI counterparts (Metzler, Eddleman, Tranor, & Cregger; 1989).

Other research over PSI within the physical education settings investigated particular components of the model, similar to investigations of the PSI model mentioned above. Leech (2011) investigated the use of flexible pacing and self-pacing in a college level instructional physical activity program (IPAP) golf course. Flexible pacing strategies included the use of instructor-recommended deadlines and student-set deadlines. Leech (2011) investigated the impact of these flexible pacing strategies on student achievement measures, course completion and withdrawal rates, and students' pacing rates. A total of 71 students participated in the study. A pretest was used to create subgroups based on skill level. Results of the study revealed that flexible pacing was advantageous for increasing the pacing rates, completion rates and student perception of the course for lower- and moderate-skilled students. Furthermore, student attitude surveys indicated that the flexible pacing strategies scored significantly greater than self-pacing on several measures including effectiveness at improving student golf ability and overall rating for the course.

Another study by Leech (2010) investigated the use of course workbooks associated with Metzler's (2001) *Personalized Sports System of Instruction* (PSIS). PSIS is a modified version of the PSI model that uses all of the original features set forth by Keller but without the use of proctors. PSIS course workbooks were developed primarily for use in college level instructional physical activity program (IPAP) but also included grade level modification suggestions that allows for use in other PE settings, including middle- and high-school settings (Metzler, 2005a). Leech (2010) investigated student perceptions of the workbook and found that students believed the workbooks were easy to use, interesting, and convenient.

Another investigation into the components of PSI within PE settings was completed by Cregger (1994). In the study, Cregger (1994) explored the effects of using a variety of presentation formats on student performance in a college-level bowling course. Presentation formats included the use of text, text and graphics, and text, graphics, and animation. Student performance was measured by students' ability to covert spares using a novice spare conversion system. Student perception data revealed that students exposed to the text, graphic, and animation presentation format perceived that they had gained greater knowledge and skill than the other groups. Analysis of the pre- and post-test data revealed no statistically significant differences within the groups based on students' ability to convert the spares.

Other research over PSI investigated methods of ensuring effective implementation of the model in PE courses. It should be noted here that much of the research completed over PSI within PE courses does not involve the use of proctors.

While this means that not all of Keller's original features are included, the flexibility of the model updates by Fox (2004) along with the litany of verification studies (Cregger & Metzler, 1992; Hannon, Holt, & Hatton, 2008; Woods, 2007) suggest that PSI can still take place without the use of proctors. Cregger and Metzler (1992) were the first to review implementation criteria for PSI in PE through their work with college-level volleyball courses. In their research, Cregger and Metzler (1992) found fourteen data sources scattered throughout four different areas could be used to determine the authenticity of PSI. These four areas included PSI course management, teacher and student processes, student progress, and student ratings of selected PSI features. Results of their study concluded that the confirmation criteria set forth by the Cregger and Metzler (1992) could be used to verify PSI implementation. The study also indicated that PSI teachers spent less than 1% of their time managing the class, less than 1% of the time lecturing, while also producing high rates of task-related feedback (0.78 per minute). The study also found that high rates of student progression and performance occurred daily, with nearly 3% of the overall tasks completed each day and 96.9% of all students completing all of the coursework by the end of the course. Along with this students' rating of PSI's effectiveness for increasing skills and knowledge were taken using a Likert scale and PSI was found to be better than average. The authors concluded that PSI was a viable alternative to conventional teaching methods for PE and that future researchers should consider investigating PSI in middle and high school settings.

Research conducted by Woods (2007) furthered the research base on the verification of effective implementation of the PSI model, focusing on the middle school level. In the study 149 middle-school PE students at a school participated in a 7-week PSI physical activity program geared towards improving students' health-related fitness components. These components included muscular strength and endurance, body composition, flexibility, and cardiorespiratory fitness. Woods (2007) used Metzler's (2005a) benchmarks to verify implementation. Results of the study revealed that PSI could be effectively implemented in a middle school PE course using the verification benchmarks developed by Metzler (2005a). Furthermore, this study was the first PSI research study involving middle-school PE students.

Within high school PE courses, research is limited, as much of the research on PSI has been completed in college-level settings (Metzler, 2005a). Work completed by Hannon, Holt, and Hatton (2008) furthered the research base on effective implementation of the PSI model. In their study, entitled *Personalized system of instruction: Teaching health-related fitness content in high school physical education*, the authors investigated the effect of PSI on student learning in a high school weight training course. In the study, 26 students enrolled in a high school weight training course were taught a post-rehabilitation unit over a three week period, using the PSI model. Data was collected using a Likert scale, audio-visual equipment, and student and teacher observations. Observational data was coded independently by two trained graduate students as well as the researchers. The researchers found that 93.4% of students met or exceeded performance criteria. Along with this, the researchers also found a high rate of feedback, a high rate of teacher cues and guidance provided by the teacher, and low amounts of management time for teachers, paralleling the results found by Cregger and Metzler (1992). The researchers concluded that based on the confirmation criteria developed by Cregger and Metzler (1992) that a PSI model could be successfully implemented in a PE weight training course with a degree of success for students.

While the above studies suggest that PSI can be effectively implemented in a variety of physical education settings and school levels, each of these studies vary drastically with respect to the content that is supposed to be learned by students. Furthermore, only Woods (2007) directly investigated PSI within a course focused on the psychomotor aspects of physical education. This raises the question of whether or not PSI can be effectively implemented in some PE courses and on some school levels more effectively than others. Along with this, other than one study completed by Pritchard, Penix, Colquitt, and McCollum (2012), recent research has not been conducted on the effectiveness of PSI in fitness-oriented PE courses (Pritchard, et al., 2012) and no research has been completed over this topic within a high school setting. Therein lies a tremendous gap in PSI's research within PE settings. In the study, Pritchard et al. (2012) investigated PSI's possible impact on student learning and achievement in a weight training course. Researchers used the Fitnessgram assessment and a fifty question knowledge test as a pre- and post-test assessment to measure the effectiveness of PSI in a fifteen-week beginning university weight training course. The Fitnessgram assessment included the progressive aerobic cardiovascular endurance run (PACER) test, back-saver sit and reach test, trunk lift test, push test, and percentage body fat test. The fifty question knowledge test (McGee & Farrow, 1987) was designed to assess overall weight training knowledge. Participants included 17 male and 5 female university students with an age range from 18 years to 48 years, (M = 20.77, SD = 6.24). A paired-samples t test with a Bonferoni correction was used to compare pre- and post-test scores. Researchers found a statistically significant difference in the pre- and post-test scores for the curl-up test, push-up test, percentage body fat test, and knowledge test. There was no statistically significant difference found between the pre- and post-test scores on the PACER, back-saver sit and reach, or the trunk lift tests. Researchers concluded that the PSI model was effective in raising achievement.

Summary

Over the last several decades, high school physical education instruction has gone through many changes. Many of these changes have been a byproduct of legislation such as the *No Child Left Behind Act* (NCLB), the Georgia SHAPE act, as well as the Race to the Top (RT3) federal initiative. These changes include the development of national and state standards for learning, standardized assessments of these standards, a raising of accountability for results of these assessments, as well as increased class sizes, and cuts in funding. Given these tremendous pressures on physical education teachers, many have begun looking for methods that research suggests can support student learning and achievement on standardized tests. One method that is being investigated is the use of gender-grouping in physical education.

Gender-grouping in physical education, as one might expect, can be somewhat controversial. The practice was in use until the Title IX legislation of 1972, which made the practice illegal until recent provisions provided by NCLB has allowed for its use once again. A summary of research over the practice in physical education yields a variety of conflicting results and centers around student perceptions and whether there exists equity between all-male and all-female course. Very little research exists on possible methods of improving student achievement in single-gender physical education courses that have a focus in the psychomotor domain and no research exists over methods to improve student Fitnessgram assessment scores within the same setting. Therefore, there exists a gap in the research as it pertains to possible methods of increasing student achievement on the Fitnessgram assessment within single-gender physical education courses that have a focus in the psychomotor domain and more research is needed. This study attempted to fill this gap in the research by investigating the possible effects of the PSI instructional model within single-gender, all-male high school physical education courses on student achievement as measured by the Fitnessgram assessment.

Another method that is being investigated is the use of personalized system of instruction (PSI) instructional model. The personalized system of instruction (PSI) instructional model is a student-centered instructional model that was developed by Keller (1968) for use in an introductory psychology course of over 300 students. Components of the original model include (1) go-at-your-own-pace, (2) unit perfection requirements, (3) the use of lectures and demonstrations as motivating factors rather than the dispersion of knowledge, (4) a focus on written word in teacher-student communication, and (5) the use of proctors (Keller, 1968). Shortly after the development of the model, the use of PSI-based courses increased dramatically over a variety of subject areas and a tremendous amount of research was completed on the model. Unfortunately, criticisms over the effectiveness of the model and the closing of the Center for Personalized Instruction due to administrative differences and funding issues (Sherman, 1992) led the instructional model to lose much of its momentum and by the end of the 1970s the number of PSI courses taught and the number of research studies over the model dropped dramatically.

Research over the PSI model is generally organized into three categories. These categories include research into the components of the model, research into effective implementation of the model, and results-based research. Most of the research that exists omits the use of proctors (Cregger, 1994; Cregger & Metzler, 1992, Hannon et al., 2008, Leech, 2010, Leech, 2011, Metzler, 2001, Tousignant, 1983, Woods, 2007) and verification of the implementation of PSI suggests this is not a cause for concern (Cregger et al., 1992). Component based research investigates the components set forth by Keller (1968) and much of this research is heavily focused on either updating these components of Keller's original plan for the 21st century, or student pacing, though there exists some research on each of Keller's original components. A summary of this line of research indicates that PSI is a flexible instructional model and that students have a greater chance of learning and of completing course objectives in a timely fashion if the instructor uses a variety of techniques, such as behavioral contracts (Brooke & Ruthven, 1984), teaching students time-management skills (Keenan, Bono, & Hursh, 1978), using a bonus point system for completing tasks early (Eppler & Ironsmith, 2004; Reidel, Harney, LaFief, & Finch, 1976; Semb, Conyers, Spencer, & Sanchez-Sosa, 1975), using student set deadlines (Roberts & Semb, 1989; Roberts & Semb, 1990), and consequences for students not meeting deadlines (Miller, Weaver, & Semb, 1974).

The second line of PSI research focuses on effectively implementing the PSI model. This line of research is primarily focused on setting criteria for implementing the model as well as the development of implementation criteria for the purpose of confirming the use of the model within the classroom. A result of this line of research is a set of specific confirmation criteria that ensures the effective implementation of the model for college-, high-, and middle school- levels.

The last line of PSI research, results-based research investigates PSI's effectiveness in raising student achievement. This line of research consists mostly of comparative research between PSI and other more traditional instructional models. Much of this research is dated at least twenty years old, with much of the research being completed while the Center for Personalized Instruction was still operational (Leech, 2011). A summary of this line of research suggests that the results of PSI is at least as effective and in many cases more effective than other instructional models in a variety of educational settings.

Within the field of physical education, PSI appears to well suited (Metzler, 2005a). Unfortunately, research over PSI in physical education settings is limited with much of the research pertaining to college-level courses (Metzler, 2005a). Furthermore, research on the effectiveness of PSI in fitness-oriented courses, such as weight training, is particularly sparse, with only one study having been completed (Pritchard, et al., 2012) in recent years. Furthermore, no research has been completed over single-gender fitness-oriented courses using PSI, and a gap in the research exists. More specifically, there exists a gap in PSI research as it pertains to high school single-gender physical education courses that have a focus in the psychomotor domain. Moreover, while the Fitnessgram has become the most frequently used physical fitness assessment used by educators (Keating & Silverman, 2004), little research has been done on the possible effectiveness of PSI in increasing high school student achievement on the Fitnessgram after controlling for prior achievement on the Fitnessgram. This lack of research is a byproduct of the lack of overall research over PSI in high school settings and constitutes a gap in PSI research. Furthermore, the foundation of PSI is built upon theoretical framework of the Social Cognitive Theory (SCT), which posits that human behavior is a product of both psychological and environmental factors (Denler, Wolter, & Benson, 2014). With this lack of research within high school single-gender physical education courses with a focus in the psychomotor domain, there exists a gap in the research pertaining to SCT within these settings.

In summary, there exist several gaps in the research that this study wishes to help to address. These gaps include the following: (1) research on social cognitive theory within single-gender, all male high school physical education courses that have a focus in the psychomotor domain, (2) possible methods of improving student achievement in single-gender, all-male physical education courses that have a focus in the psychomotor domain, and (3) the possible effect a *personalized system of instruction* instructional model has on student achievement as measured by the Fitnessgram assessment in single-gender, all-male physical education courses that have a focus in the psychomotor domain. This study will seek to address the gap in social cognitive theory within single-gender, all-male high school physical education course that have a focus in the psychomotor domain by investigating the possible impact of changes in the classroom environment from a teacher-centered, non-PSI instructional model, to a student-centered, PSI instructional model, has on student achievement as measured by the Fitnessgram assessment, after controlling for prior student achievement. This study will address possible methods of improving student achievement as measured by the Fitnessgram assessment in single-gender, all-male physical education courses that have a focus in the psychomotor domain by investigating the possible effects that an instructional model, PSI, has on student achievement as measured by the Fitnessgram assessment, after controlling for prior achievement. Finally, this study will seek to address the gap in the research on the possible effect PSI has on student achievement as measured by the Fitnessgram assessment in single-gender, all-male physical education courses that have a focus in the psychomotor domain by comparing student achievement on the Fitnessgram assessment of students who received the PSI model to students who did not receive the PSI model, controlling for prior achievement on the Fitnessgram assessment.

CHAPER 3: METHODOLOGY

The purpose of this causal comparative study was to explore the possible impact of a personalized system of instruction on student achievement on the state-mandated Fitnessgram assessments for high school physical education students within a large, urban public school system. This chapter describes the design of the study, the research questions and null hypotheses, the participants and setting, instrumentation, procedures, and how the data will be analyzed.

Design

The study was designed to be causal comparative in nature. This design was chosen for several reasons. First, there was no random selection of participants into experimental and control groups. According to Ary (2006), this eliminates all of the experimental designs as possible design options. Secondly, the independent variable being investigated, namely the type of instructional model used in the physical education classroom a student is enrolled in is not controlled by the researcher. Type of physical education class could not be controlled for because students and their guardians had the ability to choose which course to sign up for during registration based on the overall needs of the student and the restrictions due to other courses taken. According to Ary (2006), when there is no control over the independent variable, taken with the lack of randomization of participants, both experimental and quasi-experimental designs cannot be used. Further investigation into the nature of the research questions will show why the causal comparative design is most appropriate of the designs left available to the researcher.

In research question one, found in the research questions section, the purpose was to investigate a possible cause and effect relationship between the type of instructional model used in the physical education classroom environment and student performance on the progressive aerobic cardiovascular endurance run (PACER) scores on the state mandated Fitnessgram assessment, after controlling for previous achievement on the PACER. Similarly, question two attempted to investigate a possible cause and effect relationship between the type of instructional model used in the physical education classroom environment and student performance on the 90 degree push-up assessment, after controlling for previous achievement on the 90 degree push-up assessment. In a same manner, research question three investigated a possible cause and effect relationship between the independent variable, type of instructional model used in the physical education classroom environment and student performance the curl-up portion of the state mandated Fitnessgram assessment. According to Ary (2006, p. 356) in instances where the before mentioned lack of randomization of participants, the control of the independent variable is not possible, and the research questions seek to investigate a possible cause and effect relationship between the independent variable and the dependent variables, then a causal comparative design is most appropriate. Further precedence for choosing a causal comparative design can be seen in similarly structured studies performed by Floate (2011), Carroll (1998), and Riordan & Noyce (2001).

Research Questions

This causal comparative study was designed to answer several questions. These questions include the following:

RQ1: Is there a statistically significant difference between the progressive aerobic cardiovascular endurance run (PACER) scores on the state mandated Fitnessgram assessment for high school all-male physical education weight training students who have been taught using the personalized system of instruction instructional model and those who have not been taught using the personalized system of instruction instructional model, after adjusting for prior achievement on the PACER test?

RQ2: Is there a statistically significant difference between the 90 degree push-up scores on the state mandated Fitnessgram assessment for high school all-male physical education weight training students who have been taught using the personalized system of instruction instructional model and those who have not been taught using the personalized system of instruction instructional model, after adjusting for prior achievement on the 90 degree push-up test?

RQ3: Is there a statistically significant difference between curl-up scores on the state mandated Fitnessgram assessment for high school all-male physical education weight training students who have been taught using the personalized system of instruction instructional model and those who have not been taught using the personalized system of instruction instructional model, after adjusting for prior achievement on the curl-up assessment?

Null Hypotheses

Below is a description of the null hypotheses associated with each of the above research questions. Each hypothesis is presented as a null hypothesis and include the following:

$H_0$1: There will be no statistically significant difference in achievement for high school all-male physical education weight training students who have been taught using the personalized system of instruction instructional model and those students who have not been taught using the personalized system of instruction instructional model as measured by the Fitnessgram progressive aerobic cardiovascular endurance run (PACER) scores, after adjusting for differences in students' prior PACER achievement.

$H_0$2: There will be no statistically significant difference in achievement for high school all-male physical education weight training students who have been taught using the personalized system of instruction instructional model and those students who have not been taught using the personalized system of instruction instructional model as measured by the Fitnessgram 90 degree push-up scores, after adjusting for differences in students' prior 90 degree push-up achievement.

$H_0$3: There will be no statistically significant difference in achievement for high school all-male physical education weight training students who have been taught using the personalized system of instruction instructional model and those students who have not been taught using the personalized system of instruction instructional model as measured by the Fitnessgram curl-up scores, after adjusting for differences in students' prior curl-up achievement.

Participants and Setting

Below is a description of the participants and the setting for this study.

Participants

The defined population for this study included all single-gender, all-male high school weight training students. The population identified for this study consisted of high school students enrolled in an all-male weight training course in a large, urban public school system in northeast Georgia. High school students enrolled in an all-male weight training course was chosen for this study due to the popularity of all-male weight training courses in high schools in the Georgia.

Suggestions for an appropriate sample size for a causal comparative study of this nature that used a one-way analysis of covariance (ANCOVA) to confirm or reject the null hypotheses are controversial. Gay & Airasian (2003) suggest a minimum sample size of 30 participants for a causal comparative design while Olejnik (1984) suggests the minimum sample size for non-experimental uses of ANCOVA to be influenced by the initial differences between the groups. Reviewing literature, ANCOVA studies in non-experimental settings vary in sample size. Baxter, Woodward, & Olson (2001) had a total of 205 students, with 104 being in the experimental group and 101 in the control group. Baxter et al. (2001) did not clarify subgroup sample sizes. Carroll's (1998) ANCOVA study consisted of 185 total students, with 76 in being in the experimental group and 109 in the control group. Robinson's (2008) doctoral study had as few as 33 total students, with group sizes of nine and 24. For studies of this nature Gall, Gall, & Borg (2007) suggest a significance level of (α=.05). A power level of (.8) (Cohen, 1988) was used. Effect size was measured using Cohen's *d* and a medium effect size of (*d*=.5) suggested by Cohen (1977) was used in this study and is consistent with observed effect sizes on the PSI model (Pascarella & Terenzini, 1991). Given the above parameters the XLSTAT program (1995) suggested a minimum sample size of 32 per group, or 64 total participants.

There were several steps involved in drawing the sample participants from the above identified population. The experimental group of participants consisted of high school students who met the following criteria:

- The student is enrolled in an all-male weight training course for the 2014-2015 school year, and
- The student's class was taught using the personalized system of instruction (PSI) instructional model.

For the purposes of this study, a student's weight training course was identified as all-male if all the participants in the course were male. The student's class was identified as being taught using the PSI model if the teacher survey and an investigation of teacher lesson plans, pacing guides, and student materials confirmed the exclusive use of the PSI model for the course in which the student is enrolled in. An example of the teacher survey can be found in Appendix C. The criteria set forth by Cregger & Metzler (1992) and adapted by Hannon et al (2008) to confirm the use of PSI in a high school physical education courses was used during the investigation of lesson plans. Criteria include (1) self-pacing, (2) mastery-based learning, (3) teacher acting as motivator, and (4) emphasis placed on written word (p.25-26).

The control group consisted of high school students who met the following criteria:

- The student is enrolled in an all-male weight training course for the 2014-2015 school year, and
- The student's class was taught not using the personalized system of instruction instructional model.

For the purposes of this study, a student's weight training course was identified as all-male if all the participants in the course were male. The student's class was identified as not being taught using the personalized system of instruction instructional model if the teacher survey and an investigation of teacher lesson plans, pacing guides, and student materials confirmed the non-use of the PSI model.

Sampling Procedures. Convenience sampling was used in this study. This sampling method was chosen because of the accessibility of the sample to the researcher (Ary et al, 2006) and is consistent with many studies of PSI (Cregger, 1994; Hannon et al, 2008; Metzler, 1984; Pritchard et al, 2012). Both the experimental and control groups were drawn from a large, urban school where the researcher works.

There were several steps involved in drawing the sample for this study. This process was organized into the following steps: (1) gain approval for the study, (2) identify the teacher groups, (3) construct the experimental group, (4) construct the control group, and (5) finalize the sample and collect data. Below is an explanation of each step and a summary of the steps involved can be found in Appendix A.

Gaining approval for the study was the first step in drawing the sample. Consistent with the school district's policy, the researcher submitted a research proposal to the researcher's principal using the local school research request form. Following approval from the principal, the researcher faxed this form to the county office. The researcher then submitted the IRB packet and gained approval to perform the research. IRB approval forms can be found in Appendix F and G.

The next step was to identify the teacher groups. The researcher obtained from the physical education department chair a list of teachers who teach all-male weight training courses and these teachers were administered the teacher survey. The purpose of the teacher survey was to identify those teachers who were using the PSI instructional model. An example of the teacher survey can be found in Appendix C.

Next, the experimental and control teacher groups were constructed. The teachers who were identified through the teacher survey as using the PSI model were contacted and lesson plans, pacing guides, and student workbooks were collected. These materials were reviewed to confirm the use of the PSI instructional model as well as for consistency by the teacher in using the weight training curriculum that is AKS, and Fitnessgram aligned. Teachers who were confirmed to be using the PSI instructional model and the weight training curriculum comprised the experimental teacher group. Teachers who were not confirmed to be using the PSI instructional model or who were found to not be using the AKS and Fitnessgram aligned weight training curriculum were eliminated from the experimental teacher group.

Following this, the control teacher group was constructed. Using the teacher survey, the researcher identified those teachers who taught all-male weight training courses but who were not using the PSI instructional model. Next, lesson plans and pacing guides were collected. The researcher then reviewed these materials to confirm the non-use of the PSI instructional model as well as to confirm the use of the AKS and Fitnessgram-aligned weight training curriculum. Teachers found to not be using the PSI instructional model and who were using the weight training curriculum comprised the control teacher group. Teachers found to be using the PSI model or who were not using the weight training curriculum was eliminated from the control teacher group.

The final step in drawing the sample was to finalize the experimental and control groups and to collect the data. After identifying the experimental and control teacher groups, the researcher collected de-identified rosters from each group and compared the sample sizes. In the instance of unequal sample sizes, Tabachnick and Fidell (2013) suggest equalizing sample sizes by random deletion for studies of this nature (p.220). The following method for equalizing sample sizes by random deletion was used: (1) assign each student within the larger group a three digit number, (2) use a random number table to construct a sample of equal size to the other group. After assuring equal sample sizes, de-identified rosters were finalized and pre and posttest Fitnessgram data was collected. The data collected was archival data of the participants and the researcher had no interaction with the participants. The data was then imported into an excel spread sheet.

Setting

This study took take place in a large, urban public school system in northeast Georgia. The school system is composed of 77 elementary schools, 25 middle schools, and 19 high schools. Of these schools, 37 elementary, ten middle, and eight high schools have been designated as Title I schools.

According to GCPS (2012), there are nearly 164, 000 students in the school system with the following demographical distribution:

- American Indian: 0.4%
- African American: 27.9%
- Asian American: 10.3%
- Hispanic: 24.7%
- Multiracial: 3.9%
- White: 32.8 %

Of these students, (14.7%) have been identified as English Language Learners, (10.6%) as Special Education students, (12.6%) as Gifted, and (49.9%) as Free or Reduced Lunch students. Furthermore, the average teacher within the school system holds a master's degree or higher advanced degree and has twelve years of teaching experience (GCPS, 2010). In 2010, the school system was selected as a top functioning urban school system in the United States and was awarded the Broad Prize. For the 2013 school year, the system will spend on average $7,392 per student (GCPS, 2012). This site was chosen because the researcher is an employee of the system.

This study was conducted at one school within the above school system. The school day consists of four 94 minute blocks on Monday, Wednesday, and Friday with four 84 minute blocks on Tuesday and Thursday with an added advisement period during those days following first block. Each class used in this study was one semester in length and had between 35-55 students per class. Data was collected for fall and spring semesters for analysis. The pretest and posttest Fitnessgram assessments were administered in a whole-class setting and proctored by at least one Fitnessgram-trained teacher. The assessments were completed in successive order over two days and data was collected by the teacher at the completion of each assessment. The assessments were administered in a high school gymnasium.

Both the treatment and control settings for this study consisted of a high school weight room, an outdoor field, and a traditional high school classroom. The weight room consisted of fifteen workout stations, each with a multi-purpose rack, an adjustable bench, two to three barbells, and a wooden platform. Dumbbells ranging in weight from ten pounds to ninety-five pounds, kettle bells ranging in weights from twenty pounds to thirty-five pounds, dot drill mats, neck wraps, a set of plyometric jump boxes, and a variety of curl bars were also available for use during the class. A computer workstation was available to students in the weight room. The outdoor field is approximately one hundred and forty yards by sixty yards, is flat, and was used primarily for warming exercises and conditioning. The traditional classroom used in this study consisted of a student computer station, an overhead projector capable of displaying digital information, and several workstation tables used for small group and individualized instruction. A wireless internet network was available for student use in each of the above settings, and students were allowed to bring devices with internet connectivity such as smartphones and tablets to class daily.

The treatment setting for this study was all-male high school weight training courses in the school that used the personalized system of instruction (PSI) model. The focus in these classrooms was on teaching the school system's academic knowledge and skills

(AKS) for weight training using the PSI model. Teachers using the PSI model used a variety of tools to reach their curricular goals. These tools included a written student manual for each unit taught, videos that could be viewed on the student workstation, overhead projector or student smart device, along with the weight training equipment mentioned above. Instructional methods included brief whole class demonstrations and discussions followed by student led, student self-paced small group and individualized practice.

The control setting for this study was comprised of all-male high school weight training courses in the school that did not use the PSI model. Similar to the treatment setting, teachers in these classrooms used the system's AKS to drive instruction but did not use the PSI instructional model or materials. These teachers used videos that were viewed from an overhead projector, a whistle and classroom timer to set the teacher-led instructional pace, along with the weight training equipment mentioned above to meet curricular goals.

Both the control and treatment groups used the same curricular units but differed in instructional models. Table 8 provides an outline of the curricular units used in both settings, along with the associated academic knowledge and skill (AKS) codes, and Fitnessgram assessment connections for construct validity purposes.

Table 8
Weight Training Curricular Units, AKS Codes, and Fitnessgram Assessment Connections for PSI and Non-PSI Courses

Units	AKS Codes	Fitnessgram
0- Fitnessgram pretest	-	-
1 – Introduction to Weight Training, Major Muscle Groups, and Record Keeping Protocols	PEWT_A2009-1 PEWT_A2009-2 PEWT_A2009-3 PEWT_A2009-5 PEWT_A2009-7	Muscular strength, endurance, and flexibility

	PEWT_A2009-8 PEWT_A2009-9 PEWT_C2009-12 PEWT_C2009-13 PEWT_C2009-14 PEWT_D2009-15	
2 – Introduction to core lifts	PEWT_A2009-1 PEWT_A2009-4 PEWT_A2009-6 PEWT_A2009-7 PEWT_A2010-1 PEWT_B2009-10 PEWT_C2009-11 PEWT_C2009-13	Aerobic capacity, muscular strength, endurance, and flexibility
3 – Introduction to the plyometric movements	PEWT_A2009-1 PEWT_A2009-4 PEWT_A2009-6 PEWT_A2009-7 PEWT_A2010-1 PEWT_B2009-10 PEWT_C2009-11 PEWT_C2009-13	Aerobic capacity, muscular strength, endurance, and flexibility
4 – The 3x3 Block	PEWT_A2009-1 PEWT_A2009-4 PEWT_A2009-6 PEWT_A2009-7 PEWT_A2010-1 PEWT_B2009-10 PEWT_C2009-11 PEWT_C2009-13	Aerobic capacity, muscular strength, endurance, and flexibility
5- The 5x5 Block	PEWT_A2009-1 PEWT_A2009-4 PEWT_A2009-6 PEWT_A2009-7 PEWT_A2010-1 PEWT_B2009-10 PEWT_C2009-11 PEWT_C2009-13	Aerobic capacity, muscular strength, endurance, and flexibility
6 – The 5-4-3-2-1 Block and establishing baseline records	PEWT_A2009-1 PEWT_A2009-4 PEWT_A2009-6 PEWT_A2009-7 PEWT_A2010-1 PEWT_B2009-10 PEWT_C2009-11 PEWT_C2009-13	Aerobic capacity, muscular strength, endurance, and flexibility
7 – The 10-8-6 Block	PEWT_A2009-1	Aerobic capacity, muscular

	PEWT_A2009-4 PEWT_A2009-6 PEWT_A2009-7 PEWT_A2010-1 PEWT_B2009-10 PEWT_C2009-11 PEWT_C2009-13	strength, endurance, and flexibility
8 – Record breaking the 3x3 Block	PEWT_A2009-1 PEWT_A2009-4 PEWT_A2009-6 PEWT_A2009-7 PEWT_A2010-1 PEWT_B2009-10 PEWT_C2009-11 PEWT_C2009-13	Aerobic capacity, muscular strength, endurance, and flexibility
9 – Record breaking the 5x5 Block	PEWT_A2009-1 PEWT_A2009-4 PEWT_A2009-6 PEWT_A2009-7 PEWT_A2010-1 PEWT_B2009-10 PEWT_C2009-11 PEWT_C2009-13	Aerobic capacity, muscular strength, endurance, and flexibility
10 – Record breaking the 5-4-3-2-1 Block	PEWT_A2009-1 PEWT_A2009-4 PEWT_A2009-6 PEWT_A2009-7 PEWT_A2010-1 PEWT_B2009-10 PEWT_C2009-11 PEWT_C2009-13	Aerobic capacity, muscular strength, endurance, and flexibility
11 – Record breaking the 10-8-6 Block	PEWT_A2009-1 PEWT_A2009-4 PEWT_A2009-6 PEWT_A2009-7 PEWT_A2010-1 PEWT_B2009-10 PEWT_C2009-11 PEWT_C2009-13	Aerobic capacity, muscular strength, endurance, and flexibility

Note: a complete list of AKS can be found in Appendix D

While the experimental and control groups used the same curricular units and had similar materials available for use, the classroom settings between the two groups varied greatly. Teachers in the treatment group used the PSI instructional model. This instructional model is a student-centered and student-driven model with the following characteristics:

- student self-pacing, that is students move through learning modules at their own pace rather than at the teachers prescribed pace,
- mastery learning, that is a student must meet mastery performance criteria before moving on to the next learning modules,
- teacher acting as a motivator rather than as a time manager, pace-keeper, or as the primary source of knowledge,
- emphasis on written word rather than on the teacher acting as the sole source of knowledge, (Keller & Sherman, 1974; Leech, 2011; Metzler, 2005)

These classrooms can be categorized as having low lecture, demonstration, and teacher management time, high rates of teacher-student interactions, a high percentage of the class time used for practice and mastery learning of the learning objectives, and a student manual that structures the learning for students (Cregger & Metzler, 1992; Metzler, 2005).

The control group, while using the same curricular units, varied significantly from the treatment group setting. The control group setting for this study used an instructional model that is teacher-focused, teacher-paced, and teacher-led. These classrooms had a high percentage of class time used for practice and mastery learning but were teacher-driven, with students beginning and finishing learning tasks on the teacher's prompting, which is often times a whistle. Furthermore, teachers in these classrooms spent a high percentage of their time in a managerial role, rather than in student support. Further differences between the treatment and control settings can be seen in Table 9 below, which summarizes the two instructional models along seven teaching and learning features established by Metzler (2005).

Table 9
Summary of Teaching and Learning Features for the Treatment and Control Settings

Teaching and Learning Feature	Treatment Setting	Control Setting
Content Selection	Teacher maintains complete control of content decisions	Teacher maintains complete control of content decisions
Managerial Control	Teacher determines and students assume responsibility for implementing	Teacher determines and maintains control
Task Presentation	Teacher plans and students receive task information via written word (student workbook) or multi-media avenue	Teacher plans and controls all tasks
Engagement Patterns	Students practice independently of the teacher	Teacher decides which patterns will be used for each learning task
Instructional Interaction	Teacher is available for individual support and tutoring as needed	Most instructional interactions are initiated and controlled by the teacher
Pacing	Students determine the pace and progression through learning tasks	Teacher maintains control of student practice pacing
Task Progression	Student determines when to move on to the next learning task based on mastery criteria	Teacher determines when students will move on to the next learning task

Note: Portions of this table are based on Metzler (2005).

Instrumentation

The dependent variable in this study was the student's post-test health-related physical fitness level. Health-related physical fitness in this study was operationally defined as a measurement of aerobic capacity, muscular strength, endurance, and flexibility (Kinetics, 2014). The instrument used to measure these aspects of physical fitness is the Fitnessgram. The Fitnessgram is criterion referenced "fitness assessment and reporting program for youth, first developed in 1982 by The Cooper Institute"

(Kinetics, 2014). The assessment is used nationally by over 50,000 schools (Kinetics, 2014) and is endorsed by The President's Council on Fitness, Sports, and Nutrition and used in the Presidential Youth Fitness Program (The President's Council on Fitness, Sports, and Nutrition, 2014). The Fitnessgram was chosen for this study for several reasons including the following:

- it was the SPG for physical education in Georgia,
- it was the assessment chosen by the Georgia Department of Education to be used to be in compliance with the Georgia SHAPE Act,
- it has been shown to be appropriate for use in kindergarten up to adults aged 30 (Kinetics, 2014),
- it is one of the most frequently used physical fitness assessments used by educators (Keating & Silverman, 2004),
- each Fitnessgram assessment has been shown to be a valid and reliable assessment (Plowman & Meredith, 2014), and
- large scale Fitnessgram assessments administered by teachers yield reliable and valid data (Morrow, Martin, & Jackson; 2010).

The Fitnessgram measures five areas of health-related physical fitness that include: aerobic capacity, body composition, muscular strength, endurance, and flexibility. For the purposes of this study, all of the above health-related physical fitness areas except body composition will be explored. Below is a description of each of the Fitnessgram assessments used in this study along with validity and reliability information.

PACER. The progressive aerobic cardiovascular endurance run, or PACER, is one of three assessments in the Fitnessgram that measure aerobic capacity. Of the three assessments, the PACER was chosen in this study because it is the assessment mandated by the Georgia Department of Education (Kinetics, 2010) and because it is the standard assessment used by the school at which the research study took place. According to Plowman & Meredith (2014), the PACER is a progressive exercise assessment that closely simulates the treadmill test used in laboratories to measure VO_2 max levels for aerobic capacity. According to the Georgia Department of Education (2014), the assessment is set up in the following manner: within a high school gymnasium, a 20-meter horizontal space is measured off by a teacher, with a taped line separating the starting and stopping places. Students begin behind one line and an audiotape provided by the Fitnessgram is played. This audio CD explains to students how the test is set up and provides the running pace for students, defined by a beeping sound and music. A student runs from one line to the other line. A student who gets to one side before the beep must wait until the next beep before running to the other side. A student is eliminated from the assessment if the student fails to reach the other line before the beep on consecutive trips. Following elimination, the teacher and student record the lap number of elimination on the student sheet and the teacher recording sheet. Examples of these documents can be found in Appendix E. Participant scores on the PACER assessment range between one and 300.

The concurrent validity between the PACER assessment and the laboratory measured VO_2 max tests have been reviewed in many studies (Barnett, Chan, & Bruce, 1993; Leger & Gadoury, 1989; Leger & Lambert, 1982; Mahar et al., 2006, 2011; Mercier, Gadoury, & Lambert, 1988; Paliczka, Nichols, & Boreham, 1987; Ramsbottom, Brewer, & Williams, 1988). Table 10, summarizes the source, sample size, and validity coefficients from the research compiled by Plowman & Meredith (2014).

Table 10

Concurrent Validity of the PACER Assessment in Children and Adolescents

Source	Sample	Validity Coefficients
Barnett et al. (1993)	27 males, 28 females	.82, .85, .72
Boreham et al. (1990)	23 Males	.64
	18 females	.90
	23 males, 18 females	.87
Mahar et al. (2006)	135 males & females	.65
Mahar et al. (2011)	174 males & females	.75
Matsuzaka et al. (2004)	132 males & females	.74
Ruiz et al. (2008)	193 males & females	.76

Table note: summarized from Plowman & Meredith (2014).

PACER score reliability has also been widely studied. According to Plowman & Meredith (2014), reliability coefficients for the PACER assessments are consistently high. PACER score reliability is most often discussed in the form of interclass reliability and intraclass reliability. Beets and Pitetti (2006) sampled 123 males and 62 females, ages 13-18, and found the intraclass reliability to be .68 and .64. Dinschel (1994) sampled 57 males and 44 females in fourth and fifth grade and found the intraclass reliability to .84. Similar studies performed by Liu et al. (1992) and Mahar et al. (1997) found intraclass reliabilities of .93 and .90. Leger et al. (1988) sampled 139 males and females, ages six to 16, and found an interclass reliability to be .89.

90 Degree push-up. The 90 degree push-up assessment is one of three options available to educators in measuring upper body muscular strength and endurance. The 90 degree push-up was chosen because it is recommended by the Fitnessgram (Kinetics, 2014), it is the standard measurement chosen by the Georgia Department of Education, and because it is the measurement used in the school where the research took place. 90 degree push-up scores are whole numbers, ranging from zero to 99.

The objective of this assessment is to perform as many push-ups as possible at a pace of one push-up every three seconds that is set by an audio CD that is provided by the

Fitnessgram distributors (Kinetics, 2014). Students continue to perform push-ups until two form corrections are made by either the teacher or a partner. Form corrections can be made if the participant stops to rest, does not maintain the rhythmic pace set forth by the audio CD, does not achieve a 90 degree angle with the elbows on a repetition, does not maintain the correct body position with the back, or does not extend the arms fully (Kinetics, 2014).

Reliability and validity for the 90 degree push-up test, like the PACER assessment above, is quite extensive. Studies of the reliability of the 90 degree push-up test on elementary students have ranged from .64 to .99 (Saint Romain & Mahar, 2001; Tomson, 1992; Zorn, 1992), though the reliability coefficient was not directly defined. Another study performed by McManis, Baumgartner, and West (2000) sampled elementary, high school, and college students and found the intraclass stability reliability coefficients for elementary and high school students to range between .50 and .86. Lubans et al. (2011) found the intraclass stability reliability coefficients for ninth grade boys and girls to be .90 and .93 respectively. Studies on the validity coefficients for the 90 degree push-up tests have yielded a validity coefficient of .70 (Pate et al., 1993; Rutherford & Corbin, 1993), though the validity coefficient used was not defined.

Curl-up. The curl-up is the Fitnessgram recommended assessment for abdominal strength and endurance, and is the standard measurement chosen by the Georgia Department of Education (Kinetics, 2014). Scores for the curl-up are whole numbers ranging from zero to 75. To perform the curl-up, a student lies on their back on a mat, legs bent, feet flat on the floor, arms straight and flat on the mat, and with fingers stretched out. On the direction of an audio CD students curl up to the edge of the mat and then back down to the starting position at a pace of 20 repetitions per minute. Students continue at the pace provided by the CD until they can no longer perform the curl-up or until their form is corrected a second time. Corrections to form include stopping to rest, not performing the curl-up in a rhythmic fashion, not curling up far enough, or not returning to the starting position in between curl-ups (Georgia Department of Education, 2014).

Several studies have investigated the reliability and validity of the curl-up assessment. According to Plowman & Meredith (2014), test-retest reliability coefficients of .89 and .86 have been observed while single trial reliabilities of .80 for boys and .75 for girls have been observed. Other studies have tested the reliability of the curl-up assessment and have found similar results (Morrow, Martin, & Jackson, 2010). With regard to validity, the curl-up possesses both content and construct validity based on anatomical, biomechanical, and electromyography analyses (Axler & McGill, 1997; McGill, Kropf, & Steffen, 1998; Mutoh, Mori, Nakamura, & Miyashita, 1981; Noble, 1981).

The control variable in this study was the student's pre-test health-related physical fitness level. Health-related physical fitness in this study is operationally defined as a measure of aerobic capacity, muscular strength, endurance, and flexibility (Kinetics, 2014). The instrument used to measure these aspects of physical fitness is the Fitnessgram assessment. A description of each of the Fitnessgram assessments along with validity and reliability information can be found above.

The independent variable in this study was the type of instructional model used in the physical education course. The primary instrument used to collect this information was the teacher survey. The survey was based on the standards established by Cregger & Metzler (1992) that have been used in studies similar to this one to both ensure the effective implementation of the PSI model and to confirm the use of the PSI model in high school and college physical education classrooms (Cregger & Metzler, 1994; Hannon, Holt, & Hatton, 2008; Leech, 2011; Metzler, 2005). An example of the teacher survey can be found in Appendix C. Teacher interviews and a review of lesson plans, pacing guides, and student materials were also be used to confirm the instructional model.

Procedures

There were several steps involved in completing the study. Consistent with school district policies, the researcher submitted an abbreviated research proposal to the district office using the Local School Research Request form and signed by the principal of the school. This form was faxed into the Research and Evaluation office and no further district approval was needed. Following approval, the researcher submitted the proper IRB packet and gained approval to perform the research. Next the researcher began the preliminary steps for constructing the experimental and control groups by first identifying the teacher groups. The researcher contacted the physical education department head and obtained from him a list of teachers who were currently teach all-male weight training courses. Following this, the teachers on the listed were administered the teacher survey. The purpose of the teacher survey was to identify those teachers who were using the PSI instructional model in their all-male weight training course. See Appendix C for an example of the teacher survey.

Once the teacher surveys was completed, the teachers who were identified by the teacher survey as using the PSI model were contacted and lesson plans, pacing guides, and student workbooks were collected. These materials were reviewed to confirm the use of the PSI instructional model using the criteria established by Cregger and Metzler

(1992) as well as to confirm the use of the AKS and Fitnessgram-aligned weight training curriculum. Teachers who are confirmed to be using the PSI model as well as the weight training curriculum comprised the experimental teacher group. Teachers who are found to not be using the PSI model or who are not using the aligned weight training curriculum were eliminated from the experimental teacher group.

Next the control teacher group was constructed. Using the teacher survey, the researcher identified teachers who were not using the PSI model in their all-male weight training courses and lesson plans and pacing guides were collected. The researcher reviewed these materials to confirm the non-use of PSI as well as to confirm the use of the AKS and Fitnessgram-aligned weight training curriculum. Teachers confirmed to be using the weight training curriculum and who were not using the PSI model comprised the control teacher group. Teachers who were found to not be using the weight training curriculum or who were found to be using the PSI model were eliminated from the control teacher group.

Following the construction of the experimental and control teacher groups, the researcher collected de-identified rosters from each of the teacher groups and compared sample sizes. In the instance of unequal sample sizes, Tabachnick and Fidell (2013) suggest equalizing sample sizes by random deletion for studies of this nature (p. 220). The following method for equalizing sample sizes by random deletion was used: (1) assign each student within the larger group a three digit number, (2) use a random number table to construct a sample of equal size to the other group. After assuring equal sample sizes, de-identified rosters were finalized and pre and posttest Fitnessgram data was collected from teachers and imported into an excel document. The researcher then recoded each student into the experimental and control groups. Next, the assumption tests associated with each research question were performed and an analysis of covariance (ANCOVA) was used to reject or fail to reject the null hypotheses associated with each research question. Information on these procedures can be found below in the data analysis section of this manuscript.

Data Analysis

The purpose of this chapter is to develop the framework for which each of the research questions and associated null hypotheses were analyzed. Below is a description of how each research question and associated null hypothesis were statistically analyzed. Unless otherwise noted, statistical analyses will be conducted using SPSS 19.

Research Question One and the associated Null Hypothesis

The following is research question one:

RQ1: Is there a statistically significant difference between the progressive aerobic cardiovascular endurance run (PACER) scores on the state mandated Fitnessgram assessment for high school all-male physical education weight training students who have been taught using the personalized system of instruction instructional model and those who have not been taught using the personalized system of instruction instructional model, after adjusting for prior achievement on the PACER test?

The following is the null hypothesis associated with research question one:

$H_0$1: There will be no statistically significant difference in achievement for high school all-male physical education weight training students who have been taught using the personalized system of instruction instructional model and those students who have not been taught using the personalized system of instruction instructional model as measured by the Fitnessgram progressive aerobic cardiovascular endurance run (PACER) scores, after adjusting for differences in students' prior PACER achievement.

The null hypothesis seeks to find whether or not there is a statistically significant difference between the means of the experimental and control groups after group means have been adjusted for the covariate, prior PACER achievement. A one-way analysis of covariance (ANCOVA) was used to analyze this null hypothesis. ANCOVA was used for this question and null hypothesis because it allows for a comparison of the means of two groups after the means have been adjusted for the effects of a covariate (Ary, 2006; Horn, 2008; Tabachnick & Fidell, 2013). Further precedence for the use of this analysis tool, given the nature of the research question and hypothesis can be found in studies by Baxter, Woodward, and Olson (2001), Carroll (1998), Riordan and Noyce (2001), Robinson (2008), and Woodward and Baxter (1997).

There are several assumptions associated with ANCOVA that were tested. The Kolmogorov-Smirnov test was used to test for normality with a significance level more than .05 indicating normality can be assumed (Garson, 2012; Mordkoff, 2011; Robinson, 2008; Szapkiw, 2010). Scatter plots and the Pearson product-moment correlation coefficient was used to check for linearity between the CV and the DV (Horn, 2008; Szapkiw, 2010). Levene's Test for Equality of Error Variances was used to test for homogeneity of the variance with an F statistic with a significance value greater than .05 indicating equal variances can be assumed (Robinson, 2008; Szapkiw, 2010; Tabachnick & Fidell, 2013). A *F* test on the interaction of the independent variable with the covariate was used to check for the homogeneity of regression with non-significant results implying that the assumption is tenable (Horn, 2008; Tabachnick & Fidell, 2013). Suggestions for an appropriate sample size for non-experimental ANCOVA are controversial. Gay & Airasian (2003) suggest a minimum sample size of 30 participants for a causal comparative design while Olejnik (1984) suggests the minimum sample size for non-experimental uses of ANCOVA to be influenced by the initial differences between the groups. Reviewing literature, ANCOVA studies in non-experimental settings vary in sample size. Baxter, Woodward, & Olson (2001) had a total of 205 students, with 104 being in the experimental group and 101 in the control group. Baxter et al. (2001) did not clarify subgroup sample sizes. Carroll's (1998) ANCOVA study consisted of 185 total students, with 76 in being in the experimental group and 109 in the control group. Robinson's (2008) doctoral study had as few as 33 total students, with group sizes of nine and 24. For studies of this nature Gall, Gall, & Borg (2007) suggest a significance level of (α=.05). A power level of (.8) (Cohen, 1988) was used. Effect size was measured using Cohen's *d* and a medium effect size of (*d*=.5) suggested by Cohen (1977) was used in this study and is consistent with observed effect sizes on the PSI model (Pascarella & Terenzini, 1991). Given the above parameters the XLSTAT program (1995) suggested a minimum sample size of 32 per group, or 64 total participants. Descriptive statistics, (M, SD), for pretest and posttest, the adjusted M, SD for the pretest, the number (N), the number per cell (*n*), and the degrees of freedom were reported and can be found in the following chapter. The effect size was reported using the partial eta squared and was interpreted using Cohen's (1988) conventions such that .01 was considered a small effect size, .06 a moderate effect size, and .14 was considered a large effect size (p. 284 – 287). The *F* ratio was calculated along with the critical value given the above parameters and the null hypothesis was either rejected or not rejected (Tabachnick & Fidell, 2013).

Research Question Two and the associated Null Hypothesis

The following is research question two:

RQ2: Is there a statistically significant difference between the 90 degree push-up scores on the state mandated Fitnessgram assessment for high school all-male physical education weight training students who have been taught using the personalized system of instruction instructional model and those who have not been taught using the personalized system of instruction instructional model, after adjusting for prior achievement on the 90 degree push-up test?

The following is the null hypothesis associated with research question two:

H$_0$2: There will be no statistically significant difference in achievement for high school all-male physical education weight training students who have been taught using the personalized system of instruction instructional model and those students who have not been taught using the personalized system of instruction instructional model as measured by the Fitnessgram 90 degree push-up scores, after adjusting for differences in students' prior 90 degree push-up achievement.

The null hypothesis seeks to find whether or not there is a statistically significant difference between the means of the experimental and control groups after group means have been adjusted for the covariate, prior 90 degree push-up achievement. A one-way analysis of covariance (ANCOVA) was used to analyze this null hypothesis. ANCOVA was used for this question and null hypothesis because it allows for a comparison of the means of two groups after the means have been adjusted for the effects of a covariate (Ary, 2006; Horn, 2008; Tabachnick & Fidell, 2013). Further precedence for the use of this analysis tool, given the nature of the research question and hypothesis can be found in studies by Baxter, et al. (2001), Carroll (1998), Riordan and Noyce (2001), Robinson (2008), and Woodward and Baxter (1997).

There are several assumptions associated with ANCOVA that were tested. The Kolmogorov-Smirnov test was used to test for normality with a significance level more than .05 indicating normality can be assumed (Garson, 2012; Mordkoff, 2011; Robinson, 2008; Szapkiw, 2010). Scatter plots and the Pearson product-moment correlation coefficient were used to check for linearity between the CV and the DV (Horn, 2008; Szapkiw, 2010). Levene's Test for Equality of Error Variances was used to test for homogeneity of the variance with an F statistic with a significance value greater than .05 indicating equal variances can be assumed (Robinson, 2008; Szapkiw, 2010; Tabachnick & Fidell, 2013). A *F* test on the interaction of the independent variable with the covariate was used to check for the homogeneity of regression with non-significant results implying that the assumption is tenable (Horn, 2008; Tabachnick & Fidell, 2013). As previously stated, suggestions for an appropriate sample size for non-experimental ANCOVA are controversial. Gay & Airasian (2003) suggest a minimum sample size of 30 participants for a causal comparative design while Olejnik (1984) suggests the minimum sample size for non-experimental uses of ANCOVA to be influenced by the initial differences between the groups. A review literature over similar studies that used ANCOVA reveal sample sizes of varying sizes (Baxter et al., 2001; Carrol, 1998; Robinson, 2008). For studies of this nature Gall et al. (2007) suggest a significance level of (α=.05). A power level of (.8) (Cohen, 1988) was used. Effect size was measured using Cohen's *d* and a medium effect size of (d=.5) suggested by Cohen (1977) was used in this study and is consistent with observed effect sizes on the PSI model (Pascarella & Terenzini, 1991). Given the above parameters, the XLSTAT program (1995) suggests a minimum sample size of 32 per group, or 64 total participants. Descriptive statistics, (M, SD), for pretest and posttest, the adjusted M, SD for the pretest, the number (N), the number per cell (*n*), and the degrees of freedom were reported in the following chapter. The effect size was reported using the partial eta squared and will be interpreted using Cohen's (1988) conventions such that .01 was considered a small effect size, .06 a moderate effect size, and .14 was considered a large effect size (p. 284 – 287). The *F* ratio was calculated along with the critical value given the above parameters and the null hypothesis was either rejected or not rejected (Tabachnick & Fidell, 2013).

Research Question Three and the associated Null Hypothesis

The following is research question three:

RQ3: Is there a statistically significant difference between curl-up scores on the state mandated Fitnessgram assessment for high school all-male physical education weight training students who have been taught using the personalized system of instruction instructional model and those who have not been taught using the personalized system of instruction instructional model, after adjusting for prior achievement on the curl-up assessment?

The following is the null hypothesis associated with research question three:

$H_0$3: There will be no statistically significant difference in achievement for high school all-male physical education weight training students who have been taught using the personalized system of instruction instructional model and those students who have not been taught using the personalized system of instruction instructional model as measured by the Fitnessgram curl-up scores, after adjusting for differences in students' prior curl-up achievement.

The null hypothesis seeks to find whether or not there is a statistically significant difference between the means of the experimental and control groups after group means have been adjusted for the covariate, prior curl-up achievement. A one-way analysis of covariance (ANCOVA) was used to analyze this null hypothesis. ANCOVA was used for this question and null hypothesis because it allows for a comparison of the means of two groups after the means have been adjusted for the effects of a covariate (Ary, 2006; Horn, 2008; Tabachnick & Fidell, 2013). Further precedence for the use of this analysis tool, given the nature of the research question and hypothesis can be found in studies by Baxter et al. (2001), Carroll (1998), Riordan and Noyce (2001), Robinson (2008), and Woodward and Baxter (1997).

There were several assumptions associated with ANCOVA that were tested. The Kolmogorov-Smirnov test was used to test for normality with a significance level more than .05 indicating normality can be assumed (Garson, 2012; Mordkoff, 2011; Robinson, 2008; Szapkiw, 2010). Scatter plots and the Pearson product-moment correlation coefficient were used to check for linearity between the CV and the DV (Horn, 2008; Szapkiw, 2010). Levene's Test for Equality of Error Variances was used to test for homogeneity of the variance with an F statistic with a significance value greater than .05 indicating equal variances can be assumed (Robinson, 2008; Szapkiw, 2010; Tabachnick & Fidell, 2013). A *F* test on the interaction of the independent variable with the covariate was used to check for the homogeneity of regression with non-significant results implying that the assumption is tenable (Horn, 2008; Tabachnick & Fidell, 2013). As mentioned above, suggestions for an appropriate sample size for non-experimental ANCOVA are controversial and reviewing literature on similar studies using ANCOVA presents a wide range of possible sample sizes (Baxter, et al., 2001; Carrol, 1998; Robinson, 2008). Gall, et al. (2007) suggest a significance level (α=.05). A power level of (.8) (Cohen, 1988) will be used. Effect size was measured using Cohen's *d* and a medium effect size of (d=.5) suggested by Cohen (1977) was used in this study and is consistent with observed effect sizes on the PSI model (Pascarella & Terenzini, 1991). Given the above parameters the XLSTAT program (1995) suggests a minimum sample size of 32 per group, or 64 total participants. Descriptive statistics, (M, SD), for pretest and posttest, the adjusted M, SD for the pretest, the number (N), the number per cell (*n*), and the degrees of freedom were reported. The effect size was reported using the partial eta squared and was interpreted using Cohen's (1988) conventions such that .01 was considered a small effect size, .06 a moderate effect size, and .14 was considered a large effect size (p. 284 – 287). The *F* ratio was calculated along with the critical value given the above parameters and the null hypothesis was either rejected or not rejected.(Tabachnick & Fidell, 2013).

CHAPER 4: FINDINGS

The purpose as previously stated for this study was to investigate the possible impact of the *personalized system of instruction* (PSI) instructional model on student achievement on the Fitnessgram assessments. This was accomplished by comparing the achievement of students who were taught using the PSI instructional model to students who were not taught using the PSI model on the various Fitnessgram assessments, controlling for previous achievement. De-identified, archival Fitnessgram data was collected from four teachers. Of the four teachers, two teachers were identified as using the PSI model while the other two teachers were identified as not using the PSI model. The teacher survey was used to confirm the use or non-use of the PSI instructional model. Below is a description of the research questions and hypotheses, descriptive statistics of the data, and the results.

Research Questions

This study is designed to answer the following questions:

RQ1: Is there a statistically significant difference between the progressive aerobic cardiovascular endurance run (PACER) scores on the state mandated Fitnessgram assessment for high school all-male physical education weight training students who have been taught using the personalized system of instruction instructional model and those who have not been taught using the personalized system of instruction instructional model, after adjusting for prior achievement on the PACER test?

RQ2: Is there a statistically significant difference between the 90 degree push-up scores on the state mandated Fitnessgram assessment for high school all-male physical education weight training students who have been taught using the personalized system of instruction instructional model and those who have not been taught using the personalized system of instruction instructional model, after adjusting for prior achievement on the 90 degree push-up test?

RQ3: Is there a statistically significant difference between curl-up scores on the state mandated Fitnessgram assessment for high school all-male physical education weight training students who have been taught using the personalized system of instruction instructional model and those who have not been taught using the personalized system of instruction instructional model, after adjusting for prior achievement on the curl-up assessment?

Hypotheses

Below is a description of the null hypotheses associated with each of the above research questions. Each hypothesis is presented as a null hypothesis. The null hypotheses include the following:

$H_0$1: There will be no statistically significant difference in achievement for high school all-male physical education weight training students who have been taught using the personalized system of instruction instructional model and those students who have not been taught using the personalized system of instruction instructional model as measured by the Fitnessgram progressive aerobic cardiovascular endurance run (PACER) scores, after adjusting for differences in students' prior PACER achievement.

$H_0$2: There will be no statistically significant difference in achievement for high school all-male physical education weight training students who have been taught using the personalized system of instruction instructional model and those students who have not been taught using the personalized system of instruction instructional model as measured by the Fitnessgram 90 degree push-up scores, after adjusting for differences in students' prior 90 degree push-up achievement.

$H_0$3: There will be no statistically significant difference in achievement for high school all-male physical education weight training students who have been taught using the personalized system of instruction instructional model and those students who have not been taught using the personalized system of instruction instructional model as measured by the Fitnessgram curl-up scores, after adjusting for differences in students' prior curl-up achievement.

Descriptive Statistics

The analysis of the data began with an investigation of the descriptive statistics of the two groups. The experimental group was comprised of students who were, for the 2013-2014 school year, in an all-male high school weight training course that was taught using the *personalized system of instruction* (PSI) instructional model. The control group was comprised of students who were, for the 2013-2014 school year, in an all-male high school weight training course that was taught without the use of the PSI model.

Both the experimental and the control groups had 103 students. At the time of collection, the experimental group was drawn from two teachers within the school and the control group was drawn from two teachers. There were a total of 105 students that qualified for the experimental group and 103 for the control group. Using the procedures outlined in chapter three, the experimental group was trimmed so that equal sized groups could be analyzed (Tabachnick & Fidell, 2013). There were two other teachers whose students may have qualified to be in the study but teacher participation was voluntary and the teachers chose not to participate.

Following collection, the data was imported into SPSS 19 and descriptive statistics were created. Table 11 summarizes these findings for both the experimental and control groups. Figure 1 compares the means of the pretest and posttest for the experimental group and the control group for the PACER, ninety-degree push-up, and the curl-up tests. An investigation of Table 11 and Figure 1 reveals that the posttest scores were higher for both groups on every portion of the Fitnessgram assessment. Along with this, the experimental group had a higher mean than the control group on every test except for the curl-up pretest, while the spread of the control group data, as measured by the standard deviation, was greater on every test than the experimental group.

Table # 11
Fitnessgram Descriptive Statistics Based on Group

Test	*M*	*SD*
Experimental Group (*Personalized System of Instruction*) (*n* = 103)		
Pacer Pretest	22.88	7.47
Pacer Posttest	29.46	8.86
Push-up Pretest	18.37	5.16
Push-up Posttest	21.48	5.59
Curl-up Pretest	25.15	8.09
Curl-up Posttest	31.95	10.14
Control Group (*Non-Personalized System of Instruction*) (*n* = 103)		
Pacer Pretest	21.78	8.78
Pacer Posttest	27.05	9.08
Push-up Pretest	16.89	6.61
Push-up Posttest	19.74	6.59
Curl-up Pretest	26.17	10.27
Curl-up Posttest	31.15	10.37

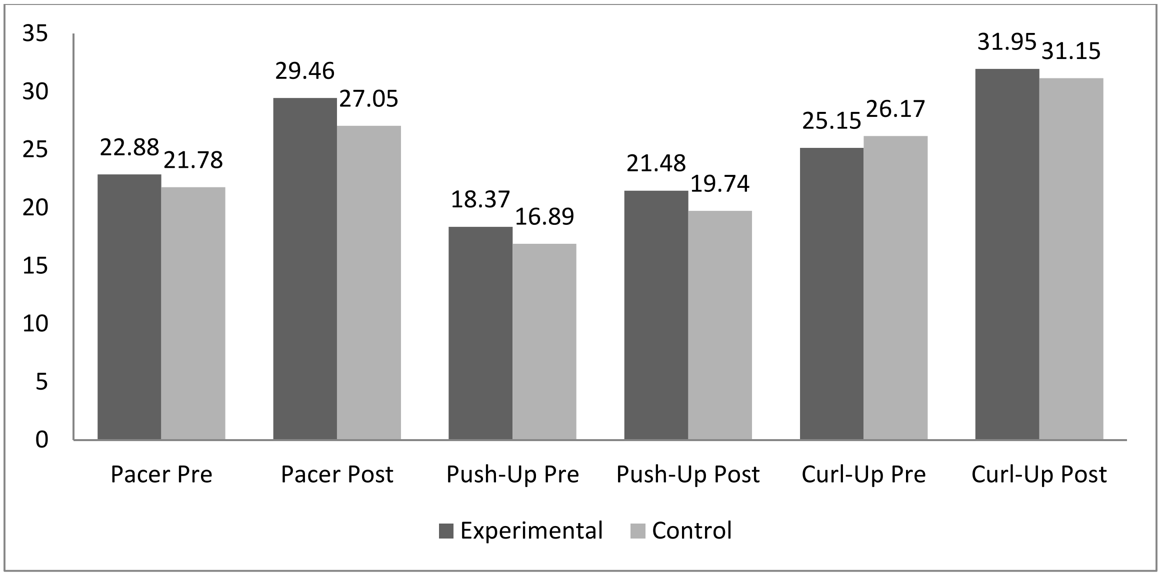

Figure 1. Comparison of the means of the experimental group (*Personalized System of Instruction*) and the control group (*Non-Personalized System of Instruction*) for the pre and posttest PACER, push-up, and curl-up.

Following an analysis of the descriptive statistics, each of the null hypotheses associated with the research questions was investigated. The results of this investigation are below.

Results

Null Hypothesis One: PACER Scores Comparison

Research question one, found above seeks to investigate whether or not there is a statistically significant difference in the PACER scores for students taught using the PSI instructional model and those students taught without the PSI instructional model after controlling for prior PACER achievement. The related null hypothesis, found above, states there will be no statistically significant difference between the PACER scores after controlling for prior PACER achievement. For the purposes of this question, the independent variable was defined as the type of instructional model that students received and was categorically defined as PSI (experimental group) or non-PSI (control group).

The dependent variable was defined as students' PACER posttest scores and the covariate was defined as students' PACER pretest scores. An analysis of covariance (ANCOVA) was used to investigate this research question.

There were several assumptions associated with ANCOVA that were analyzed prior to the development of the ANCOVA model. Normality was checked using boxplots, frequency histograms, and residual plots. Linearity between the covariate, PACER pretest scores, and the dependent variable, PACER posttest scores, was checked using scatter plots and the Pearson product-moment correlation coefficient ($|r| = .85$), which denotes a very strong linear relationship allowed for the assumption to be assumed. Levene's Test for Equality of Variance was used to test for the homogeneity of variance ($p = .27 > .05$), indicating that equal variances could be assumed (Robinson, 2008; Szapkiw, 2010; Tabachnick & Fidell, 2013). A *F* test on the interaction of the independent variable, instructional model used, with the covariate, PACER pretest scores, was used to check for the homogeneity of regression ($p = .35 > .05$) with the non-significant results implying the assumption was tenable (Horn, 2008; Tabachnick et al., 2013).

Following the assumption tests, a one-way ANCOVA model was created. Table 13 summarizes these findings.

Table 13
Comparison of Posttest PACER Scores Controlling for the Pretest (N=206)

Source	SS	*df*	*MS*	*F*	*p*	Partial Eta Squared
Full Model	12131.581	2	6065.790	268.427	.001	.726
Pacer Pretest	11833.018	1	11833.018	523.643	.001	.721
Group	96.714	1	0.811	0.036	.85	.001
Error	4587.293	203	22.598			
Total	16718.874	205				

Note: Levene's Test of Equality of Error Variances $p = .27$

Inspection of the table found a significant PACER pretest (*p* = .001). The main effect for the experimental group was not significant (*p* = .85). This combination of findings provided support to fail to reject the null hypothesis.

Null Hypothesis Two: Ninety-Degree Push-Up Scores Comparison

Research question two, found above seeks to investigate whether or not there is a statistically significant difference in the ninety-degree push-up scores for students taught using the PSI instructional model and those students taught without the PSI instructional model after controlling for prior ninety-degree push-up achievement. The related null hypothesis, found above, states there will be no statistically significant difference between the scores after controlling for prior achievement. For the purposes of this question, the independent variable was defined as the type of instructional model that students received and was categorically defined as PSI (experimental group) or non-PSI (control group). The dependent variable was defined as students' ninety-degree push-up posttest scores and the covariate was defined to students' ninety-degree push-up pretest scores. An analysis of covariance (ANCOVA) was used to investigate this research question.

There were several assumptions associated with ANCOVA that were analyzed prior to the development of the ANCOVA table. Normality was checked using boxplots, frequency histograms, and residual plots. Linearity between the covariate, push-up pretest scores, and the dependent variable, push-up posttest scores, was checked using scatter plots and the Pearson product-moment correlation coefficient ($|r| = .87$), which denotes and strong linear relationship and that the assumption can be assumed. Levene's Test for the Equality of Variance was used to test for the homogeneity of variance ($p = .21 > .05$), indicating that equal variances could be assumed (Robinson, 2008; Szapkiw, 2010; Tabachnick et al., 2013). A *F* test on the interaction of the independent variable, instructional model used, with the covariate, push-up pretest scores, was used to check for the homogeneity of regression ($p = .30 > .05$) with the non-significant results implying the assumption was tenable (Horn, 2008; Tabachnick, et al., 2013).

Following the assumption tests, a one-way ANCOVA model was created. Table 14 summarizes these findings.

Table 14
Comparison of Posttest Push-Up Scores Controlling for the Pretest (N=206)

Source	SS	*df*	*MS*	*F*	*p*	Partial Eta Squared
Full Model	5848.640	2	2924.320	309.749	.001	.753
Push Pretest	5693.101	1	5693.101	603.023	.001	.748
Group	9.059	1	9.059	0.960	.33	.005
Error	1916.511	203	9.441			
Total	7765.150	205				

Note: Levene's Test of Equality of Error Variances $p = .21$

Reviewing the Table 14, found significant push-up pretest ($p = .001$). The main effect for the experimental group was not significant ($p = .33$). This combination of findings provided support to fail to reject the null hypothesis.

Null Hypothesis Three: Curl-Up Scores Comparison

Research question three, found above seeks to investigate whether or not there is a statistically significant difference in the curl-up scores for students taught using the PSI instructional model and those students taught without the PSI instructional model after controlling for prior curl-up achievement. The related null hypothesis, found above, states there will be no statistically significant difference between the scores after controlling for prior achievement. For the purposes of this question, the independent variable was defined as the type of instructional model that students received and was categorically defined as PSI (experimental group) or non-PSI (control group). The dependent variable was defined as students' curl-up posttest scores and the covariate was defined to students' curl-up pretest scores. An ANCOVA model was used to investigate this research question.

There were several assumptions associated with ANCOVA that were analyzed prior to the development of the ANCOVA model. Normality was checked using boxplots, frequency histograms, and residual plots. Linearity between covariate, curl-up pretest scores, and the dependent variable, curl-up posttest scores, was checked using scatter plots and the Pearson product-moment correlation coefficient was found to denote a strong linear relationship ($|r| = .84$). Levene's Test for the Equality of Variance was used to check the homogeneity of variance and the non-significant results indicated that equal variances could be assumed ($p = .72 > .05$). A *F* test on the interaction of the independent variable, instructional model used, with the covariate, curl-up pretest scores, was used to check for the homogeneity of regression and the non-significant results indicated that the assumption was tenable ($p = .28 > .05$).

Following the assumption tests, an ANCOVA model was created. Table 15 summarizes these findings.

Table 15
Comparison of Posttest Curl-up Scores Controlling for the Pretest (N=206)

Source	*SS*	*df*	*MS*	*F*	*p*	Partial Eta Squared
Full Model	15404.070	2	7702.035	257.286	.001	.717
Curl-Up Pretest	15370.628	1	15370.628	513.455	.001	.717
Group	161.301	1	50.832	1.731	.19	.008
Error	6076.945	203	29.936			
Total	21481.015	205				

Note: Levene's Test of Equality of Error Variances $p = .72$

Reviewing the Table 15, found significant curl-up pretest (p = .001). The main effect for the experimental group was not significant (p = .19). This combination of findings provided support to fail to reject the null hypothesis.

In summary, this study used archival data to test the social cognitive theory that relates the instructional model received by a student to student achievement, controlling for prior student achievement for high school physical education students. The null hypothesis that is associated with research question one (differences in PACER scores) was not rejected (Table 13). The null hypothesis associated with research question two (differences in push-ups) was not rejected (Table 14). The null hypothesis that is associated with research question three (differences in curl-up scores) was not rejected (Table 15). In the final chapter, these findings will be compared to literature, conclusions and implications will be drawn, limitations of the study will be discussed, and a series of recommendations will be suggested.

CHAPER 5: DISCUSSION, CONCLUSION, AND RECOMMENDATIONS

The previous chapter presented data analyses of the research hypotheses utilizing the ANCOVA statistical procedures to examine whether or not there existed a statistically significant difference in the achievement of students on the Fitnessgram assessments based on the instructional model received, after controlling for prior student achievement on the Fitnessgram assessments. This chapter is organized into a discussion of the findings in light of a review of the current literature, followed by conclusions, implications of the findings, limitations of the study, and finally recommendations for future research.

Discussion

The purpose of this study was to test the social cognitive theory that relates the instructional model received by a student to student achievement within a high school all-male weight training course. This was done by comparing the achievement of high school weight training students who had been taught using the *personalized system of instruction* (PSI) instructional model to students who had not been taught using the PSI instructional model over the battery of Fitnessgram assessments, after controlling for prior achievement on the Fitnessgram assessments. Below is a discussion of the findings of the study in light of current research and literature over the *personalized system of instruction* within physical education. As mentioned in Chapter 2, current results-based research that investigates PSI's effectiveness in raising student achievement is limited, with much of the research having been completed at least twenty years ago while the Center for Personalized Instruction was still in operation (Leech, 2011). Furthermore, research on the effectiveness of PSI in fitness-oriented courses, such as weight training, is particularly sparse, with only one study having been completed (Pritchard, Penix, Colquitt, & McCollum; 2012) in recent years. The discussion below is organized by research question and null hypothesis. The discussion will review the findings of the current study and compare the results to the results found by Pritchard et. al (2012), the only current study of similar nature. Following this, a summary of the findings will be compared to historical findings. Finally, a summary of the findings with respect to social cognitive theory will be discussed.

Discussion of the Results of Research Question One and Null Hypothesis One

The purpose of research question one, which can be found in Chapter 3, was to investigate whether or not there was a statistically significant difference between the PACER scores on the Fitnessgram assessment for high school all-male physical education weight training students who have been taught using the PSI instructional model and those who had not been taught using the PSI instructional model, after adjusting for prior PACER achievement. The null hypothesis stated there would be no statistically significant difference between the two groups. A review of the statistical analyses of the null hypothesis resulted in not rejecting the null hypothesis ($p = .85$). More specifically, no statistically significant difference was found between the PACER scores of those students who had been taught using the PSI instructional model and those who had not been taught using the PSI instructional model, after controlling for previous PACER achievement.

Reviewing this result in the context of current results-based PSI research is limited to one study performed by Pritchard, Penix, Colquitt, and McCollum (2012). A review of these results with respect to historical PSI research will be discussed later in this chapter. Along with this, a discussion of the impact of these findings with respect to social cognitive theory will also be discussed later in this chapter.

In their study, Pritchard, et. al (2012) used the Fitnessgram assessments as well as a fifty question knowledge test as a pre- and post-test assessment to attempt to measure the effectiveness of PSI in a university-level weight training course. The Fitnessgram assessment included the PACER test, back-saver sit and reach test, trunk lift test, push-up test, and the percentage body fat test. The fifty question knowledge test (McGee & Farrow, 1987) was designed to assess overall weight training knowledge. Twenty-two students participated in the study with an age range from 18 years to 48 years. A paired-samples t-test with Bonferroni correction was used to measure for statistical significance between student pretest scores and posttest scores. The researchers found no significant difference in PACER scores. The current study supports these findings, as no statistically significant difference was found between the PACER scores of the PSI and non-PSI groups, after controlling for prior PACER achievement. It should be noted that the two studies seek to fill gaps in the research over PSI but over different populations (high school all-male weight training courses versus college-level coeducational courses). Therefore, while it is interesting to review the findings of the two studies together, findings should also be viewed independently of each other as well.

Discussion of the Results of Research Question Two and Null Hypothesis Two

Similar to research question one, the purpose behind research question two was to investigate whether or not there was a statistically significant difference between the ninety-degree push-up scores on the Fitnessgram assessment for high school all-male physical education weight training students who have been taught using the PSI instructional model and those who have not been taught using the PSI instructional model, after adjusting for prior ninety-degree push-up achievement. The null hypothesis stated there would be no statistically significant difference between the two groups of students. A review of the statistical analyses of the null hypothesis resulted in not rejecting the null hypothesis (p = .50). More specifically, no statistically significant difference was found between the ninety-degree push-up scores of those students who had been taught using the PSI instructional model and those who had not been taught using the PSI instructional model, after controlling for previous ninety-degree push-up achievement.

As was the case with the review of the results for research question one, reviewing the results of research question two in the context of current results-based PSI research is limited to the study performed by Pritchard et. al (2012). A review of these results with respect to historical research will be discussed later in this chapter. Along with this, a discussion of the impact of these findings with respect to social cognitive theory will also be discussed later in this chapter.

In their study, Pritchard, et. al (2012) used the Fitnessgram assessments as well as a fifty question knowledge test as a pre- and post-test assessment to attempt to measure the effectiveness of PSI in a university-level weight training course. The Fitnessgram assessment included the PACER test, back-saver sit and reach test, trunk lift test, push-up test, and the percentage body fat test. The fifty question knowledge test (McGee & Farrow, 1987) was designed to assess overall weight training knowledge. Twenty-two students participated in the study with an age range from 18 years to 48 years. A paired-samples t-test with Bonferroni correction was used to measure for statistical significance between student pretest scores and posttest scores. The results of their study found a statistically significant difference between push-up pretest and posttest scores. In one aspect, the current study does not support these findings, as no statistical significance was found between the ninety-degree push-up scores of those taught using the PSI model and those not taught using the PSI model, after controlling for prior student achievement on the test. In another aspect, due to the differences in the sample populations, high school all-male students versus college-level males and females with a thirty year age-range, it is in the opinion of this researcher that the two studies vary enough to make both results independent of each other and not contradictory. Each of the two studies seek to fill gaps in the research over PSI but over very different populations and educational settings, therefore the results, though interesting to viewed together, should also be viewed as independent of each other and not necessarily contradictory.

Discussion of the Results of Research Question Three and Null Hypothesis Three

The purpose of research question three, which can be found in Chapter 3, was to investigate whether or not there was a statistically significant difference between the curl-up scores on the Fitnessgram assessment for high school all-male physical education weight training students who have been taught using the PSI instructional model and those who have not been taught using the PSI instructional model, after adjusting for prior curl-up achievement. The null hypothesis stated there would be no statistically significant difference between the two groups. A review of the statistical analyses of the null hypothesis resulted in not rejecting the null hypothesis ($p = .19$). More specifically, no significant difference was found between the curl-up scores of those students who had been taught using the PSI instructional model and those who had not been taught using the PSI instructional model, after controlling for previous curl-up achievement.

As previously noted, current results-based research over the PSI model is limited to the study performed by Pritchard et. al (2012) and a discussion of these findings with respect to historical literature will be discussed later in this chapter along with a discussion of the impact of these results on social cognitive theory research. Pritchard et.

al (2012) investigated the possible impact of PSI on student pretest and posttest curl-up scores using a paired-samples t-test with Bonferroni correction. The results of their study found a statistically significant difference between the pretest and posttest curl-up scores. As before, a cursory review of the two studies finding might suggest the current study does not support these findings, as no statistically significant difference was found. On the other hand, due to the differences between the sample populations, all-male high school students versus coeducational college-level students, it is in the opinion of this researcher that the two studies and their results stand independently of each other and do not necessarily contradict each other. Each of the two studies seek to fill gaps in the research over PSI but over different populations and educational settings, therefore though interesting to view together, these findings should also be viewed as independent of each other and not necessarily contradictory.

Summary of Findings Compared with Historical Research over PSI in Physical Education

As noted in Chapter two, research over PSI in physical education (PE) courses is limited (Metzler, 2005a), even though many believe the instructional model to be well-suited to PE courses (Colquitt, Pritchard, & McCollum, 2011; Hannon, Holt, & Hatton, 2008; Metzler, 2005a, Pritchard et. al, 2012). Much of the recent research over the model investigates the various components of the instructional model, or focuses on ways of ensuring effective implementation of the model, both of which fall outside of the scope of this research study. In many ways this gap in the research is a product of an overwhelming amount of results-based research over the model that had been done in the past over a variety of subject matters. As is the case with PSI research across other educational fields, much of the research over PSI in PE is dated, having been completed in the 1970's while the Center for Personalized Instruction was still operational (Leech, 2011). One of the purposes of this study was to help to fill the gap in research as it pertains to current results-based research over the PSI instructional model within PE courses. That being said, a discussion of the findings of this study as it compares with the most common historical studies' findings is of importance and it is to that we now turn.

Reviewing the finding of the three research questions, no statistically significant difference was found between those students who were taught using the PSI instructional model and those who were taught without the use of the PSI instructional model.

Findings such as these would at first glance appear to contradict research completed by Annarino (1976), who found that PSI results to be equally or more effective than other teaching methods but the scope of the results above does not allow for conclusions to be drawn that would contradict Annarino's (1976) findings. The results of the current study suggest that there is no statistically significant difference in scores, after controlling for prior achievement which would confirm Annarino's findings that PSI results to be at least equal to other teaching methods. Along with this, Annarino's (1976) research study reported results from many independent studies that included a variety of sample populations and educational levels that allow his results and the present results to stand independently and without contradictory results.

Several studies by Metzler were also compared to the current study. These studies like that of Annarino (1976) were chosen for review because in reviewing literature, these studies were most often referred to and are considered to be a foundation of PSI research within the field of physical education. Along with this, the Metzler studies are the most recent of the historical research reviewed that is similar in nature to the present study.

In studies completed by Metzler in 1984, 1986, 1988, the researcher reported among other results, higher rates of success for PSI students across a spectrum of outcome criteria when compared to their direct instruction counterparts. To a certain degree, the non-significant results of the current study contradict the findings of Metzler (1984), Metzler (1986), and Metzler (1988), as the current finding suggest there to be no statistically significant difference between PSI and non-PSI achievement as measured by the Fitnessgram assessment, after controlling for prior achievement. Fortunately, there is enough of a difference between the current study and those performed by Metzler to allow for a common ground. First, Metzler compared the PSI instructional model to the direct instruction (DI) instructional model exclusively while the current study compared PSI and non-PSI instructional models. This difference in comparison groups allows for both studies to find and fill different gaps in the research over the PSI instructional model. Finally, Metzler (1984), Metzler (1986), and Metzler (1988) used college-level physical education students while the current study was performed with all-male high school students, once again allowing for these studies to fill different gaps in the research over the PSI model and not necessarily be on contradictory terms.

Other studies that often come up in a review of literature over the PSI instructional model include studies performed by Cregger (1994), Cregger and Metzler (1992), Fox (2004), Hannon, Holt, and Hatton (2008), Leech (2010), Mezler, Eddleman, Tranor, and Cregger (1989), Tousignant (1983), and Woods (2007). While each of these studies add to the growing literature over the PSI model, these studies investigate topics such as updating components of the PSI model or investigations into effective implementation of the model that fall outside the realm of comparison with the current study's findings.

Summary of Findings with Respect to Social Cognitive Theory

The results of this study support Bandura's Social Cognitive Theory which posits that a part of an individual's knowledge acquisition is a result of the influence of psychological and environmental factors working in a reciprocal manner (Parajes, 2009). As applied to my study, one might interpret the theory to suggest that environmental factors, such as the instructional model used in a classroom, to influence student learning and achievement because the instructional model is a form of environmental change that the theory postulates would "automatically lead to changes in the person('s)" (Boston University, 2013, para. 4) learning and behavior. Continuing on this line of reasoning, one might come to the conclusion that the non-significant results found for each of the null hypotheses contradict or at the very least, do not support the claims of the theory. Fortunate for us, this line of reasoning would be incorrect.

Reviewing the precepts of the theory, note the theory only suggests "parts" of knowledge acquisition to be a result of environmental factors. Furthermore, Banduras (1989) reminds us that these sources of causation (psychological and environmental factors) do not imply the sources to be of equal strength (p.2). In reviewing the results of the study, the effect size of the groups, as measured by the partial eta squared, ranged from .001 to .008, implying that the type of instructional model received by a student accounted for between .1% and .8% of their posttest score. This increase in achievement is by no means significant, even to the most passionate teacher looking to change a student's life one point, or repetition, at a time, but the findings do suggest that environment, or in this case instructional method used, does play a very small part in student achievement in the study.

Along with these findings, this study supports Bandura's (1986) claim that personal factors, for example biological events and genetics, as well as behavioral factors, for example a physically active lifestyle, does impact student learning and achievement. Reviewing the results of the study, the effect size, as measured by the partial eta squared, of the pretest scores ranged from .717 to .748. These findings imply that whatever combination of personal and behavior factors that make a student's prior physical fitness levels greatly influences (between 71.7% to 74.8%) Fitnessgram posttest student achievement.

Conclusion

In conclusion, the purpose of this study was to contribute to the growing body of research over social cognitive theory that relates instructional model received by a student to student achievement. This study attempted to do this by comparing the achievement of high school all-male weight training students who had been taught using the *personalized system of instruction* (PSI) instructional model to those students who had not been taught using the PSI instructional model, after controlling for prior achievement. By exploring the possible impact of PSI on high school student achievement, this study also attempted to contribute to the body of research on PSI in high school PE weight training courses in large high schools.

Following the analyses of the three research questions, there was found to be no statistically significant difference between the Fitnessgram PACER, ninety-degree push-up, and curl-up tests achievement of students taught using the PSI instructional model and those not taught using the PSI instructional model, after controlling for prior student

Fitnessgram achievement and each of the null hypotheses were not rejected. These results appear to support current results-based research on the impact of PSI on the Fitnessgram PACER test and to contradict the research on PSI's impact on the Fitnessgram ninety-degree push-up and curl-up tests but the current results-based research is limited to one study performed by Pritchard et. al (2012). Furthermore, a review of historical results-based research suggests that this study at first glance does not support the claims presented but due to the differences in sample populations and educational settings, the current study and historical findings should be viewed not only together but also independently and as attempting to fill different gaps in the research. With respect to social cognitive theory, the current study supports Banduras' (1989) claims that both psychological and environmental factors have the ability to impact learning and that these factors are not necessarily equal in impact.

The problem this study set out to address was that while many PE teachers and department chairs are considering implementing the PSI instructional model in high school PE classes in hopes of improving student achievement on the state mandated Fitnessgram assessment, there is very little current research on the effectiveness of the PSI model in high school PE classes in raising student achievement. This study added to this very limited research base and the implications of this study are discussed below.

Implications

The implications of this study are multi-faceted. As the pressures on high school physical education teachers continues to rise due to ever expanding class sizes, more accountability for results on state mandated tests, and budget cuts, physical education (PE) teachers and departments are continually searching for instructional models that research suggests supports student achievement. In this process many are turning to the *personalized system of instruction* (PSI) instructional model. Unfortunately, much of the results-based research over the model within PE settings was completed over twenty years ago, was completed on a university-level, or was conducted in non-fitness-oriented courses or class units. This gap in the research over the PSI instructional model as it pertains to results-based research within fitness-oriented high school PE courses was directly addressed in this study. Along with this, the current study contributed to the growing literature on social cognitive theory and how it may apply in high school PE settings.

Reviewing the non-significant findings of this study calls into question the effectiveness of the PSI model in improving student achievement over the current instructional model that is being used within high school all-male weight training courses and more research is needed. With the costs in investing in the PSI instructional model, both the monetary costs to purchase training materials, to update older classrooms to make them more PSI-ready, and the professional development hours in training, the implications of the findings of this study should at the very least make a physical education teacher, department chair, or other steak holder pause before a full scale adoption by the whole PE department of the PSI instructional model.

For those teachers already using the model in their high school weight training courses, those early adopters who have already gone to conferences, collaborated with colleagues, and effectively implemented the model, the findings of this study should give you hope and reinforcement in what you are doing. The non-significant results of this study should inform you that, after controlling for prior Fitnessgram achievement, there is no statistically significant difference in the achievement of your students from those of the traditional, teacher-centered, “old-school whistle blower” weight training teachers who still litter current PE departments and will often heckle you for your style of teaching. Keep doing what you are doing. Keep facilitating, encouraging, and communicating to this generation of students in their language that often includes the videos and other digital media that the PSI instructional model lends itself to. Continue to mold the culture of your classroom into a student-centered one that encourages proctors, written word, and other PSI foundational ideals. Keep on keeping on! Go back to a graduate program, do the research, and help fill the gap in current results-based PSI research!

Limitations

This study has several limiting factors. These factors include the following:

- This study was limited to one large high school, located in a large, urban school district in north-east Georgia.
- This study was limited to high school all-male weight training courses.
- This study cannot account for the quality of the teacher, which could have affected student performance on the Fitnessgram assessments.
- This study cannot account for the motivational levels of the participating students, which could have affected student performance on the Fitnessgram assessments.
- This study was causal-comparative in nature, which limits any possible cause-effect relationship findings between the independent variable, instructional model, and the dependent variable, Fitnessgram posttest scores.
- This study relied upon archival data collected from teachers who volunteered to participate in the study, making this study self-selected.

1. The researcher did not monitor any classrooms and relied upon a research-based teacher survey to determine the full implementation of the PSI instructional model.

Recommendations for Future Research

To understand the possible relationship between the instructional model used by a teacher and student performance within fitness-oriented physical education classes, this study compared the achievement of all-male high school weight training students who had been taught using the *personalized system of instruction* (PSI) instructional model to those who had not been taught using the PSI instructional model, after controlling for prior achievement. Through the development of this study, several avenues for future reach into the PSI instructional model began to become apparent. The first recommendation is based on one of the gaps in the research that this study attempted to address, namely, there is very limited current results-based research over PSI within physical education fitness-oriented courses and more is needed. As mentioned above, results-based research consists of research that compares the results of PSI to the results of another instructional model on state mandated standardized tests that teachers are accountable for. As the pressures on physical education teachers continues to rise due to ever expanding class sizes, more accountability for results on state mandated tests, and budget cuts, physical education teachers and department chairs are seeking instructional models that research suggests supports student achievement in fitness-oriented classes. Unfortunately, many teachers and department chairs are turning to a PSI instructional model that has very little current research support. Further results-based research is needed. Along with this, there are several other recommendations to be made for future research over the PSI model within physical education courses. These recommendations are focused on closing the gap within result-based research over PSI. These recommendations include the following:

- Conduct a result-based study across multiple physical education fitness-oriented courses that are restricted to all-female high school students to see if PSI impacts student achievement for those courses and students.
- Conduct a results-based study across multiple high school physical education fitness-oriented courses that would investigate the possible impact of PSI on student achievement on state mandated tests across high school grade levels: freshmen, sophomores, juniors, and seniors.
- Conduct a longitudinal study that measures the impact of PSI on student physical fitness levels, tracking students from the beginning of their freshmen year to the end of their senior year.
- Conduct a results-based study over the possible impact of PSI on high school Fitnessgram low-achievers within beginner weight training courses.
- Conduct a results-based study over the possible impact of PSI on high school Fitnessgram high-achievers within fitness-oriented physical education courses.
- Conduct a results-based study over the possible impact of PSI on middle school student achievement on the Fitnessgram assessments.
- Conduct a study similar to the current study in a different school, county, state, and/or region of the United States.

REFERENCES

Addinsoft. (1995). XLSTAT program. Retrieved from http://www.xlstat.com/en/download.html

Ary, D. (2006). *Introduction to research in education* (7th ed.). Belmont, CA: Thomson/Wadsworth.

Associates, L. P. (2007). Understanding the no child left behind act: Technology integration quick key 3 Retrieved June 2011, from http://www.learningpt.org/pdfs/qkey3.pdf

Atherton, J. (2009). Learning and Teaching: Behaviorism Retrieved June 2010, from http://www.learningandteaching.info/learning/behaviour.htm.

Au, W. (2009). Social studies, social justice: W(h)ither the social studies in high-stakes testing? *Teacher Education Quarterly, 36*(1), 43-58.

Axler, C., & McGill, S. (1997). Low back loads over a variety of abdominal exercises: Searching for the safest abdominal challenge. *Medicine and Science in Sports Exercise,* 29 (6), 804-811.

Banduras, A. (1977). Social learning theory. Englewood Cliffs, NJ: Prentice Hall.

Banduras, A. (1989). Social cognitive theory. *Annals of Child Development*, 6, 1-60.

Baxter, J., Woodward, J., & Olson, D. (2001). Effects of reform-based mathematics instruction on low achievers in five third-grade classrooms. *The Elementary School Journal,* 101(5), 529-547.

Barnett, A., Chan, L., & Bruce, I. (1993). A preliminary study of the 20-m multistage shuttle run as a predicator of peak VO2 in Hong Kong Chinese students. *Pediatric Exercise Science*, 5(1), 42-50.

Barrett, T. (2005). The effects of cooperative learning on performance of sixth grade physical education students. *Journal of Teaching in Physical Education,* 24, 88-102.

Beets, M., Pitetti, K. (2006). Criterion-referenced reliability and equivalency between the PACER and 1-mile run/walk for high school students. *Journal of Physical Activity & Health,* 3.

Blacke, C. (2012). Single-sex education versus coeducation in north Georgia public middle schools. Retrieved from digitalcommons.liberty.edu/doctoral/606.

Boston University. (2013). *Social cognitive theory.* Retrieved from http://sphweb.bumc.bu.edu/otlt/MPH-Modules/SB/SB721-Models/SB721-Models5.html

Bradley, K. (2009). An investigation of single-sex education and its impact on academic achievement, discipline referal frequency, and attendance for first and second grade public education students. Retrieved from ProQuest Dissertations and Theses database. (UMI No. 3374139)

Brehe Pixler, P. (2009). *Mild disability students and Everyday Mathematics: Serving the needs of this population.* Ed.D. 3399155, The University of Toledo, United States -- Ohio. Retrieved from http://proquest.umi.com/pqdweb?did=1978310481&Fmt=7&clientId=20655&RQT=309&VName=PQD

Brooke, R., & Ruthven, A. (1984). The effects of contingency contracting on student performance in a PSI class. *Teaching of Psychology,* 8, 87-89.

Buskist, W., Cush, D., & DeGrandpre, R. (1991). The life and times of PSI. *Journal of Behavioral Education*, 1(2), 215-234.

Carroll, W. (1998). Geometric knowledge of middle school students in a reform-based mathematics curriculum. *School Science and Mathematics*, 98(4), 188-197.

Center on Budget and Policy Priorities. (2013). Most states funding schools less than before the recession. Retrieved from http://www.cbpp.org/cms/?fa=view&id=4011

Cohen, J. (1977). *Statistical power analysis for the behavioral sciences* (Rev. ed.). New York: Academic Press.

Cohen, J. (1988). *Statistical power analysis for behavior sciences* (2nd ed.). Hillsdale: Erlbaum.

Colquitt, G., Pritchard, T., McCollum, S. (2011). The personalized system of instruction in fitness education, *Journal of Physical Education, Recreation & Dance,* 46-54.

Conard, C. (1997). Fairbanks system of instruction at Fairbanks Country Day. *Behavior and Social Issues,* 7(10), 25-29.

Cregger, R. (1994). *Effects of three presentation formats in the PSI college level bowling course*. Virginia Polytechnic Institute and State University. **Doctoral**.

Cregger, R., Metzler, M. (1992). PSI for a college physical education basic instructional program. *Educational Technology, 32,* 51-56.

Creswell, J. W. (2007). *Qualitative inquiry & research design : choosing among five approaches* (2nd ed.). Thousand Oaks: Sage Publications.

Crone-Todd, D. & Pear, J. (2001). Applications of Bloom's taxonomy to PSI. *Behavior Analyst Today,* 3, 204-210.

Denler, H., Wolters, C., & Benzon, M. (2013). Social Cognitive Theory. Retrieved from http://www.education.com/reference/article/social-cognitive-theory/

Dinschel, K. (1994). Influence of agility on the mile run and PACER tests of aerobic endurance in fourth and fifth grade school children. Unpublished master's thesis, Northern Illinois University, KeKalb, IL.

Doherty, J. Teaching styles in physical education and Mosston's spectrum. (2010). *Kentucky AHPERD*, 4-6.

DuFour, R., DuFour, R. B., & Eaker, R. E. (2008). *Revisiting professional learning communities at work : new insights for improving schools*. Bloomington: Solution Tree.

Drexel University. (2009), from http://mathforum.org/mathed/constructivism.html

Dunkin, J. (1986). Research on teaching in higher education. *Handbook on Research of Teaching*, 754-777.

Dyson, B. (2001). Cooperative learning in an elementary physical education program. *Journal of Teaching in Physical Education,* 20, 264-281.

Dyson, B., Griffin L., Hastie, P. (2004). Sport education, tactical games, and cooperative learning: Theoretical and pedagogical considerations. *Quest,* 56, 226-240.

Eppler, M. & Ironsmith, M. (2004). PSI and distance learning in a developmental psychology course. *Teaching of Psychology,* 31(2), 131-134.

Eyre, H. (2007). Keller's personalized system of instruction: Was it a fleeting fancy or is there a revival on the horizon? *The Behavior Analyst Today,* 8 (3).

FairTest.org. (2007). Gender bias in college admissions tests. Retrieved from www.fairtest.org/facts/genderbias.htm

Finn, C. (1989). *Real reform in education*. Washington, DC: Government Printing Office.

Floate, J. The effects of physical activity choice on physical fitness of northern Mississippi middle school students. *Doctoral Dissertations and Projects.* Paper 469. http://digitalcommons.liberty.edu/doctoral/469.

Fox, E. (2004). The personalized system of instruction: A flexible and effective approach to mastery learning. *Evidence-based Educational Methods,* 201-221.

Fullan, M. (2001). *Leading in a culture of change* (1st ed.). San Francisco: Jossey-Bass.

Gabbei, R. (2004). Achieving balance: Secondary physical education gender-grouping options. *The Journal of Physical Education, Recreation, & Dance, 75*(3), 33-39.

Gall, M. D., Gall, J. P., & Borg, W. R. (2007). *Educational research : an introduction* (8th ed.). Boston: Pearson/Allyn & Bacon.

Gallup, H., & Allan, R. (2003). Concerns with recent criticisms of the personalized system of instruction (PSI). Lafayette College website.

Garson, G. (2012). Testing statistical assumptions. *North Carolina Statistical Associates Publishing.*

Gay, L. R., & Airasian, P. W. (2003). *Educational research : competencies for analysis and applications* (7th ed.). Upper Saddle River, N.J.: Merrill/Prentice Hall.

Georgia Department of Education. (2008a). 2008 CRCT score interpretation guide: Grades 1 thru 8. Atlanta

Georgia Department of Education. (2008b). An assessment and accountability brief: Validity and reliability for the 2008 criterion-referenced-competency tests.

Georgia Department of Education. (2009). Talking points. Retrieved from http://archives.gadoe.org/DMGetDocument.aspx/Fitness%20Assessment%20Talking%20Points%20May%202011.pdf?

Georgia Department of Education. (2010). Curriculum, instruction, and assessment: Criterion-referenced competency tests. Retrieved from http://www.doe.k12.ga.us/ci_testing.aspx?PageReq=CI_TESTING_CRCT

Georgia Department of Education. (2011). Criterion-referenced competency tests spring 2011: School and system test coordinator's manual, from http://www.doe.k12.ga.us/DMGetDocument.aspx/CRCT%20STC%20Manual%20Spring%202011.pdf?p=6CC6799F8C1371F6C0C283BF6D703C977C5BF235AD050171A278C5C275CCA314&Type=D

Georgia Department of Education. (2012a). Enrollment by gender, race/ethnicity and grade (PK-12).

Georgia Department of Education. (2012b). Georgia student growth model. Retrieved from http://www.doe.k12.ga.us/Curriculum-Instruction-and-Assessment/Assessment/Documents/SGP%20FAQ%20122112.pdf.

Grady, M., L. (2009). *Constructivst mathematics theory: The effects of Everyday Mathematics on sixth grade achievement on the Illinois Standards Mathematics Achievement Test*. Retrieved from Pro Quest Dissertation and Thesis database database (5, 70). Dissertation Abstract International.

Graham, G. (1981). Research on teaching in physical education: A discussion with Larry Locke and Daryl Siedentop. *Journal of Teaching in Physical Education,* 1(1), 3-15.

Grant, L., & Spencer, R. (2003). The personalized system of instruction: Review and applications to distance learning. *The International Review of Research in Open and Distributed Learning*, 4(2).

Griffin, L., & Butler, J. (2005). Teaching games for understanding: Theory, research, and practice. Champaign, IL: Human Kinetics.

Gurvitch, R., & Metzler, M. (2013). Aligning learning activities with instructional models. *Journal of Physical Education, Recreation, & Dance,* 84(3), 30-37.

Gurvitch, R., Metzler, M., & Lund, J. (2008). Model-based instruction in physical education: The adoption of innovation. *Journal of Teaching in Physical Education,* 27(4), 447-589.

Gwinnett County Public Schools. (2011). About US, from http://www.gwinnett.k12.ga.us/gcps-mainweb01.nsf/pages/Home-AboutUs1~MainPage

Gwinnett County Public Schools (2014a). "Gwinnett Effectiveness Initiative." from http://publish.gwinnett.k12.ga.us/gcps/home/public/employment/gei.

Gwinnett County Public Schools (2014b). "Gwinnett County Effectiveness Initiative - Teachers."

Hamilton, L.S., Stecher, B.M., Yuan, K. (2008). Standards-based reform in the United States: History, research, and future directions. Retrieved from www.rand.org/content/dam/rand/pubs/reprints/2009/RAND_RP1384.pdf.

Hannon, J. C., Holt, B.J., Hatten, J.D. (2008). "Personalized system of instruction model: Teaching health-related fitness content in high school physical education." *Journal of Curriculum and Instruction, 2*(2), 20-32.

Hannon, J.C, & Ratliffe, T. (2005). Physical activity level in coeducational and single-gender high school physical education settings. *Journal of Teaching in Physical Education, 24,* 149-164.

Hannon, J., & Ratliffe, T. (2007). Opportunities to participate and teacher interactions in coed versus single-gender physical education settings. *Physical Educator, 64*(1), 11-20.

Hannon, J., & Williams, S. (2008). Should secondary physical education be coeducational or single-sex? *JOPERD: The Journal of Physical Education, Recreation & Dance, 79*(2), 6-56.

Hansen, G., Brothen, T., & Wambach, C. (2002). An evaluation of early alerts in PSI general psychology course. *Learning Assistance Review,* 7(1), 14-21.

Hill, G., Hannon, J., & Knowles, C. (2012). "Physical Education Teachers' and University Teacher Educators' Perceptions Regarding Coeducational versus Single Gender Physical Education." *The Physical Educator*, 69, 265-288.

Horn, R. A. (2008). Understanding analysis of covariance (ANCOVA). Retrieved from http://oak.ucc.nau.edu/rh232/courses/EPS625/Handouts/ANCOVA/Understanding%20ANCOVA.pdf

Howell, D. C. (2008). *Fundamental statistics for the behavioral sciences* (6th ed.). Belmont, CA: Thomson/Wadsworth.

Hill, R.L. (2011). The effects of single gender education on the achievement of sixth grade mathematics students. Retrieved from digitalcommons.liberty.edu/doctoral/467.

Hurst, D. (2009). The impact of reform-based elementary mathematics textbooks on students' fractional number sense. Retrieved from ProQuest Digital Dissertations. (UMI: 3363756).

Hymel, G. (1987). A literature trend analysis in mastery learning. Paper presented at the Annual Meeting of the American Educational Research Association, Washington, D.C.

Institute of Education Sciences. (n.d.). Retrieved from http://ies.ed.gov/aboutus/

Ironsmith, M., Eppler, M.A. (2007). Mastery learning benefits low-aptitude students. *Teaching Psychology, 34,* 28-31.

Issacs, A., Carroll, W., & Bell, M. (2001). A research-based curriculum: The research basis of the UCSMP everyday mathematics curriculum.

Joyce, B., Weil, M., & Calhoun, E. . (2000). *Models of teaching* (6th ed.). Boston: Allyn & Bacon.

Juker, D., McGill, S., Kropf, P., & Steffen, T. (1998). Quantitative intramuscular myoelectric activity of lumbar portions of psoas and the abdominal wall during a wide variety of tasks. *Medicien and Science in Sports and Exercise*, 30(2), 301-310.

Keller, F. (1968). *Goodbye, teacher!* Journal of Applied Behavior Analysis, 1, 79-88.

Keller, F., Sherman, J. (1974). *The Keller plan handbook.* Menlo Park, CA: W.A. Benjamin.

Kelly, M. (2014). *How budget cuts affect teachers.* Retrieved from http://712educators.about.com/od/issuesineducation/tp/teaching_budget_cuts.htm

Keenan, J., Bono S., & Hursh, D. (1978). Shaping time-management skills: Two examples in PSI. *Journal of Personalized Instruction,* 3, 46-49.

Kinetics, H. (2014). "Fitnessgram Program Overview - Teachers." from http://www.fitnessgram.net/program-overview/teacher-overview.

Knowledgebase, L. T. (2011). Constructivism, from http://www.learning-theories.com/constructivism.html

Kober, N., Chudowsky, N., & Chudowsky, V. (2008). *Has student achievement increased since 2002?: State test score trends through 2006-07*. Washington, DC: Center on Education Policy.

Koca, C. (2009). Gender interaction in coed physical education: A study in Turkey. *Adolescence,* 44(173), 165-185.

Kulik, J. (1976). PSI: A formative evaluation. *Personalized instruction in higher education: Proceedings of the the second national conference.* Washington, D.C.: Center for Personalized Instruction.

Kulik, J. & Bangert-Downs, R. (1990). Effectiveness of mastery learning programs: A meta-analysis. *Review of Educational Research,* 60, 265-299.

Kulik, J. Jaska, P., & Kulik, C. (1978). Research on component features of Keller's personalized system of instruction. *Journal of Personalized Instruction,* 3(1), 2-14.

Kulik, J., Kulik C, & Cohen, P. (1979). A meta-analysis of outcome studies of Keller's personalized system of instruction, *American Psychologist*, 34 (4), 307-318.

Lamal, P. (1984). Interest in PSI across sixteen years. *Teaching of Psychology*, 11, 237-238.

Leech, J.R. (2010). An examination of pacing rates, course completion rates, and perceptions in a college-level PSI-taught gold course.

Leech, J. R. (2011). Effects of pacing contingencies in a PSI-taught college-level golf course. Education, Florida State University. Doctor of Philosophy.

Leger, L., & Gadoury, C. (1989). Validity of the 20 m shuttle run test with 1 minute stages to predict VO2 max in adults. *Canadian Journal of Sport Sciences,* 14(1), 21-26.

Leger, L., Mercier, D., Gadoury, C., & Lambert, J. (1988). The multistage 20 meter shuttle run test for aerobic fitness. *Journal of Sports Sciences,* 6(2), 93-101.

Leger, L., & Lambert, J. (1982). A maximal multistage 20-m shuttle run test to predict VO2 max. *European Journal of Applied Physiology and Occupational Physiology,* 49(1), 1-12.

Lirgg, C.D. (1993). Effects of same-sex versus coeducational physical education on self-perceptions of middle and high school students. *Research Quarterly and Exercise Science, 64*, 324-334.

Lowry, W., & Thornburg, M. (1988). A working biography of the Keller plan (PSI). Logan, UT.

Lui, H. (2003). Development of an online course using a modified version of Keller's personalized system of instruction. Unpublished doctoral dissertation, Virginia Polytechnic Institute and State University, Blacksburg.

Magliaro, S. G., Lockee, B.B., & Burton, J.K. (2005). Direct instruction revisited: A key model for instructional technology. *Educational Technology Research & Development, 53*(4), 41-53.

Marzano, R. J. (2003). *What works in schools: Translating research into action*. Alexandria, VA: Association for Supervision and Curriculum Development.

McGraw Hill. (2012). Everyday mathematics classroom manipulatives: Grade PK-6: Grade 5 manipulatives. Retrieved from https://www.mheonline.com/program/view/1/16/233/0076020339/.

McGee, R., & Farrow, A. (1987). Test questions for physical education activities. Champaign, IL: Human Kinetics.

McKenzie, T., Prochaska, J., Sallis, J., & LaMaster, K. (2004). Coeducational and single-sex physical education in middle schools: Impact on physical activity. *Research Quarterly for Exercise and Sport, 75,* 446-449.

McManis, B., Baumgartner, T., & West, D. (2000). Objectivity and reliability of the 90 degree pushup test. *Measurement in Physical Education and Exercise Science,* 4, 57-67.

Metzler, M. (1984). Analysis of a mastery learning/Personalized system of instruction for teaching tennis. *The 1984 Olympic Scientific Congress Proceedings,* 6, 63-67.

Metzler, M. (1986). *Teaching tennis by "traditional" and PSI-based instruction.* Paper presented at the meeting of the American Educational Research Association, San Fransisco, CA.

Metzler, M. (1988). A review of research in sport pedagogy. *Journal of Teaching in Physical Education,* 8, 87-103.

Metzler, M. (2001). Mastering the basics of the personalized system of instruction system. Boston: Allyn & Bacon.

Metzler, M. (2005). *Instructional models for physical education* (2nd ed.). Scottsdale, AZ, Holcomb Hathaway.

Metzler, M., Eddleman, K., Treanor, L., & Cregger, R. (1989). Using a personalized sytem of instruction for teaching tennis. Paper presented at the Eastern Educational Research Association Annual Meeting, Suvannah, GA.

Metzler, M., Sebolt, D. (1994). *Instructor's manual for the personalized sport instruction series.* Dubuque, IA: Kendall Hunt.

Miller, L., Weaver, F., & Semb, G. (1974). A procedure for maintaining student progress in a personalized university course. *Jouranl of Personalized Instruction,* 7, 81-91.

Mordkoff, T. (2011). The assumption(s) of normality. Retrieved from http://www2.psychology.uiowa.edu/faculty/mordkoff/GradStats/part%201/I.07%20normal.pdf

Morrow, J., Martin, S., & Jackson, A. (2010). Reliability and validity of the Fitnessgram: Quality of teacher-collected health-related fitness surveillance data. *Research Quarterly for Exercise and Sport*, 81(3), 24-30.

Mosston, M. (1966). Teaching physical education. Columbus, OH: Merrill.

Mutoh, Y., Mori, T., Nakamura, Y., & Miyashita, M. (1983). The relationship between sit-up exercises and the occurance of low back pain. *Biomechanics VIII-A*, 180-185.

National Association for Sport and Physical Education. (1995). Moving into the future: National standards for physical education: A guide to content and assessment. Reston, VA.

National Center for Education Statistics. (2009). Executive summary: Achievement gaps: How black and white students in public schools perform in mathematics and reading on the National assessment of education progress. Retrieved from: http://nces.ed.gov/nationsreportcard/pubs/studies/2009455.asp

National Center for Education Statistics. (2009). Executive summary: Achievement gaps: How hispanic and white students in public schools perform in mathematics and reading on the National assessment of education progress. Retrieved from: http://nces.ed.gov/nationsreportcard/pubs/studies/2011459.asp.

National Center for Education Statistics. (2010). NAEP overview: About NAEP. Retrieved from http://nces.ed.gov/nationsreportcard/

National Center for Education Statistics. (2013a). NAEP overview. Retrieved from http://nces.ed.gov.nationsreportcard.

National Center for Education Statistics. (2013b). Cautions in interpreting NAEP results. Retrieved from http://nces.ed.gov/nationsreportcard/about/interpretresults.asp.

National Commission on Educational Excellence. (1983). A nation at risk: The imperative for education reform from http://www.ed.gov/pubs/NatAtRisk/intro.html

National Council of Teachers of Mathematics. (1989). *Curriculum and evaluation standards for school mathematics*. Reston, VA: National Council of Teachers of Mathematics.

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics*. Reston, VA: National Council of Teachers of Mathematics.

National Council of Teachers of Mathematics. (2011). Journals and Books Retrieved June, 2011, from http://www.nctm.org/standards/content.aspx?id=11608

National Institute for Direct Instruction (2014a). *Basic philosophy of direct instruction.*

National Institute for Direct Instruction (2014b). *DI vs di.*

National Science Foundation (1988). *The NSF: A brief history.* Washington, DC: National Science Foundation, Directorate for Education and Human Resources.

Nelson, K.A. (2014).A study of comparing fifth grade student achievement in mathematics in departmentalized and non-departmentalized settings. *Doctoral*

Dissertations and Projects. Paper 829. http://digitalcommons.liberty.edu/doctoral/829

Nilges, L. (1998). I thought only fairy tales had supernatural power: A radical feminist analysis of Title IX in physical education. *Journal of Teaching in Physical Education,* 17(2), 172-194.

Nixon, J., & Locke, L. (1973). Research on teaching physical education. *Second Handbook of Research on Teaching*, 1210-1242.

Noble, L. (1981). Effects of various types of situps on iEMG of the abdominal musculature. *Journal of Human Movement Studies,*7, 124-130.

Olejnik, S. F. (1984). Planning educational research: Determining the necessary sample size. *The Journal of Experimental Education, 53*(1), 40-48.

Olafson, L.L. (2002). "I hate phys. ed." : Adolescent females talk about physical education. *Physical Educator, 59*(2), 67-74.

Otten, S. (2011). Journals publishing work in mathematics education, from www.tandf.co.uk/journals/titles/10986065.asp

Painter, J. (2008). Everyday Math, from http://www.rationalamerican.com/rp.org/archives/everydaymath.html

Paliczka, V., Nichols, A., & Boreham, C. (1987). A multi-stage shuttle run as a predictor of running performance and maximal oxygen uptake in adults. *British Journal of Sports Medicine,* 21(4), 163-165.

Pascarella, E., & Terenzini, P. (1991). *How college affects students: Findings and insights from twenty years of research.* San Francisco: Jossey-Bass.

Pate, R., Ward, D. Saunders, R., Felton, G., Dishman, R., & Dowda, M. (2005). Promotion of physical activity among high-school females: A randomized controlled trial. *American Journal of Public Health,* 95(9), 1582-1587.

Pfister, G. (2005). Berkshire encyclopedia of world sport volum 1. Berkshire Publishing Group, Great Barrington, MA.

Pixler. P. B. (2009). Mild disability students and Everyday mathematics: Serving the needs of this population. Retrieved from ProQuest Digital Dissertations. (UMI 3399155).

Plowman, S., & Meredith, M. (2014). Fitnessgram/Activitygram Reference Guide. Retrieved from https://www.cooperinstitute.org/vault/2440/web/files/662.pdf.

Pritchard, T., Colquitt, G. (2006). Personal fitness: Maximizing student learning. Paper presented at the Annual Meeting of the American Alliance for Health, Physical Education, Recreation and Dance, Salt Lake City, UT.

Pritchard, T., Penix, K., Colquitt, G., McCollum, S. (2012). Effects of a weight training personalized system of instruction course on fitness level and knowledge. *The Physical Educator,* 69, 342-359.

Prize, T. B. (2010). The Broad Prize - Gwinnett County Public Schools, GA., from http://www.broadprize.org/asset/1579-tbp%202010%20gwinnett%20fact%20sheet.pdf

ProCon. (2009). Glossary of Terms, from http://2008election.procon.org/view.resource.php?resourceID=001575

PSU. (n.d.) Retrieved from http://info.psu.edu.sa/psu/maths/007%20Social%20Cog%20Theory%2001.pdf.

Ramsbottom, R., Brewer, J., & Williams, C. (1988). A progressive shuttle run test to estimate maximal oxygen uptake. *British Journal of Sports Medicine,* 22(4), 141-144.

Rand Corporation. (2008). Standards-based reform in the United States: History, research, and future directions.

Ravitich, D. (1995). *National standards in American education: A citizen's guide*. Washington, DC: Brookings.

Reys, B., Dingman, S., Sutter, A., & Teuscher, D. (2005). Development of state-level mathematics curriculum documents: Report of a survey. Retrieved from http://www.mathcurriculumcenter.org/ASSM_report.pdf

Reidel, R., Harney, B., Lafief, W., & Fince, M. (1976). The effect of time as a contingency on student performance in an individualized course. *Personalized instruction in higher education: Proceedings of the second national conference.* Washington, D.C.: Georgetown University, Center for Personalized Instruction, 128-130.

Roberts, T. (2009). Relationships between students' fitness levels and academic achievement. *Doctoral Dissertations and Projects.* Paper 235. Retrieved from http://digitalcommons.liberty.edu/doctoral/235.

Roberts, M., & Semb, G. (1989). Student selection of deadline conditions in a personalized psychology course. *Teaching Psychology,* 16, 128-130.

Roberts, M., & Semb, G. (1990). Analysis of the number of student-set deadlines in a personalized psychology course. *Teaching Psychology,* 17(3), 170-173.

Robin, A. (1974). Behavioral instruction in the classroom: A review. Stony Brook, NY: State University of New York Instructional Resources Center.

Robinson, A.H. (2008). Highly qualified teacher status and the reading achievement of students with disabilities. Retrieved from Liberty Digital Commons. Paper 112.

Rodenroth, K. (2010). A study of the relationship between physical fitness and academic performance. *Doctoral Dissertations and Projects.* Paper 358. Retrieved from http://digitalcommons.liberty.edu/doctoral/358

Riordan, J., & Noyce, P. (2001). The impact of two standards-based mathematics curricula on student achievement in Massachusetts. *Journal for Research in Mathematics Education, 32*, 368-398.

Rude, B. (1996). Chicago Math, from http://brianrude.ecom/chi-mth.htm.

Rutherford, J., & Corbin, C. (1993). Measuring upper body strength and endurance: Which test is best. *Kentucky AHPERD Journal,* 29(2), 20-24.

Sadker, M., Sadker, D., & Long, L. (1993). Gender and educational equality. *Multicultural Education: Issues and Perspectives,* 111-128.

Saint Romain, B., & Mahar, M. (2001). Norm-referenced and criterion-referenced reliability of the push-up and modified pull-up. *Measurement in Physical Education and Exercise Science,* 5(2), 67-80.

Schmitt, M. (2001). *A comparison of physical activity levels of males and females in co-educational and gender segregated sixth grade physical education classes.* Eugene, OR: Microform Publications, University of Oregon.

Schoenfeld, A. (2002). Making mathematics work for all children: Issues of standards, testing, and equity. *Educational Researcher, 31*(1), 13-25.

Schultz, G. (1998). *Kingdom education : God's plan for educating future generations*. Nashville, TN: LifeWay Press.

Semb, G., Conyer, D., Spencer, R., & Sanchez-Sosa, J. (1975). An experimental comparison of four pacing contingencies in a personalized instruction course. *Behavior Research and Technology in Higher Education.*

Sherman, J. (1982). PSI today. *The PSI Handbook: Essays on personalized instruction,* 72-78.

Sherman, J. (1992). Reflections on PSI: Good news and bad. *Journal of Applied Behavioral Analysis,* 25(1), 59-64.

Siedentop, D. (1974). How to use personalized system of instruction in college teaching. *National college Physical Education Association for Men Proceedings,* 77th annual meeting, Kansas City, 116-125.

Silberman, R. (1978). The Keller plan: A personal view. *Journal of Chemical Education,* 55(2), 97-98.

Silverman, S. (1991). Research on teaching physical education. *Research Quarterly for Exercise and Sport,* 62, 352-364.

Sinclair, C. (2000). How do adolescents perceive coeducational and same-sex physical education classes? *Journal of Physical Education, Recreation, & Dance, 71*(1), 9.

Singer, R., & Dick, W. (1974). Teaching physical education: A system approach. Boston, MA: Houghton Mifflin Co.

Sizemore, C. B. (2010). *Voices of reform: An oral history of the impact of mathematics reform on Texas public schools: 1960-2010.* Doctor of Education Dissertation, Baylor University, United States - Texas. Retrieved from http://proquest.umi.com.ezproxy.liberty.edu:2048/pqdweb?index=0&did=2260543161&SrchMode=1&sid=1&Fmt=2&VInst=PROD&VType=PQD&RQT=309&VName=PQD&TS=1329780572&clientId=20655

Slavin, R. E. (2009). *Educational psychology : theory and practice* (9th ed.). Boston: Pearson/Allyn and Bacon Publishers.

Slavin, R. E., & Lake C. (2008). Effective programs in elementary mathematics: A best evidence synthesis. *Review of Educational Research, 78*(3), 427-515.

Special Education. (2006). WordNet: A lexical database for the English language, from http://www.wordnet.princeton.edu/perl/webwn.

Svennengsen, L., Pear, J. (2011). Effects of computer-aided personalized system of instruction in developing knowledge and critical thinking in blended learning courses. *The Behavior Analyst Today,* 12(1).

Szapkiw, A. (2010). Statistics Guide.

Tabachnick, B.G. and Fidell, L.S. (2013). *Using Multivariate Statistics* (6th ed.). Boston, MA: Pearson.

Taveggia, T. (1976). Personalized instruction: A summary of comparative research, 1967-1975.

The Century Foundation. (2014). *State spending cuts weaken public schools.* Retrieved from http://tcf.org/work/education/detail/state-spending-cuts-weaken-public-schools

The President's Council on Fitness, Sports, and Nutrition. (2014). Presidential youth fitness program. Retrieved from http://www.fitness.gov/participate-in-programs/presidential-youth-fitness-program/

The University of Northern Iowa. (2011). SPSS Techniques Series: Statistics on Likert Scale Surveys, from http://www.uni.edu/its/support/article/604#chi

Tomson, L. (1992). A field trial of strength and endurance for primary grade children. Unpublished master's thesis, Arizona State University, Tempe, AZ.

Tousignant, M. (1983). PSI in PE - it works. *Journal of Physical Education, Recreation and Dance, 54,* 33-34.

U.S. Department of Education. (2002). Executive Summary: Why NCLB is important, from http://www2.ed.gov/nclb/overview/intro/execsumm.html

U.S. Department of Education. (2005). Introduction: No Child Left Behind archived information, from http://www2.ed.gov/nclb/overview/intro/index.html?exp=5

U.S. Department of Education. (2008). New race and ethnicity guidance for the collection of federal education data, from http://www2.ed.gov/policy/rschstat/guid/raceethnicity/index.html

Van Dalen, D., & Bennett, B. (1971). *A world history of physical education: Cultural, philosophical, comparative.* Englewood Cliffs.

Viadero, D. (2008). Research effort aims to bury 'nothing works' image. *Education Week.* Retrieved from http://www.edweek.org/ew/articles/2008/12/15/16whatworks.h28.html.

Wageman, J., & Bartoszuk, K. . (2004). Adequate yearly progress (AYP) and no child left behind: What could AYP mean for your child, school, and community. *EXT: NDSU Extension Service*.

Waite, R. D. (2000). *A study of the effects of Everyday Mathematics on student achievement of third-, fourth-, and fifth-grade students in a large north Texas urban school district.* Ed.D. 9992659, University of North Texas, United States -- Texas. Retrieved from http://proquest.umi.com/pqdweb?did=727816971&Fmt=7&clientId=20655&RQT=309&VName=PQD

Ward, P., & Lee, M. (2005). Peer-assisted learning in physical education: A review of theory and research. *Journal of Teaching in Physical Education,* 5, 205-225.

Wellstone, P. D. (2000). High stakes test: A harsh agenda for America's children. Retrieved from www.educationrevolution.org/paulwellstone.html

What Works Clearinghouse (2010a). *Intervention Report: Everyday Mathematics*. Institute of Education Sciences

What Works Clearinghouse (2010b). What Works Clearinghouse: Everyday Mathematics, from http://ies.ed.gov/ncee/wwc/reports/advancedss.aspxt

What Works Clearinghouse (2011). Procedures and Standards Handbook, from http://ies.ed.gov/ncee/wwc/documentsum.aspx?sid=19

What Works Clearinghouse (n.d.-a). Review Process, from http://ies.ed.gov/ncee/wwc/reviewprocess.aspx

What Works Clearinghouse (n.d.-b). WWC Frequently Asked Questions Retrieved June 2011, from http://ies.ed.gov/ncee/wwc/Document.aspx?sid=15&pid=

Wichita State University. (n.d.). *The Keller Plan.* Retrieved from http://webs.wichita.edu/depttools/depttoolsmemberfiles/belder/Question%202%20-%20Keller%20Plan.htm

Williams, M. W. (2009). *Comparision of fifth-grade students' mathematics achievement as evidenced by Georgia's criterion-referenced competency test: Traditional and departmentalized settings.* Doctorate of Education Dissertation, Liberty University, Lynchburg.

Wilson, Z. J. (2012). *The effects of single-gender classes on students' physical fitness test performances and attitudes*. Education, Liberty University. Doctor of Education.

Wittenburg, R. A., Northup, K.L., & Cottrell, L.A. (2012). "Children's aerobic fitness and academic acheivement: A longitudinal examination of students during their fifth and seventh grade years." American Journal of Public Heath **102**(12): 2303-2307.

Woods, M. (2007). Physical activity beliefs and middle school PE instruction. Unpublished manuscript. University of Illinois at Urbana-Champaign.

Woodward, J. (2009). *A study of physical fitness and academic performance levels of sixth and seventh grade students.* Doctoral Dissertations and Projects. Paper 309. Retrieved from http://digitalcommons.liberty.edu/doctoral/309

Woodward, J., & Baxter, J. (1997). The effects of an innovative approach to mathematics on academically low-achieving students in inclusive settings. *Exceptional Children, 63*(3), 373-388.

Zorn, R. (1992). Selected tests of muscular strength and endurance for children. Unpublished master's thesis, Arizona State University, Tempe, AZ.

APPENDIX A: Research Steps

Step	**Explanation**
	A. GAIN APPROVAL
A.1	Submit a research proposal to the county using the abbreviated research plan using the Local School Research Request Form and signed by the Principal. No further approval is necessary.
A.2	Chair submits the proper IRB forms.
	B. IDENTIFYING TEACHER GROUPS
B.1	Contact the physical education department head and obtain a list of teachers who teach all-male weight training courses.
B.2	Administer the teacher survey in order to identify PSI and non-PSI instructional model teachers.
	C. CONSTRUCT EXPERIMENTAL GROUP
B.1	Contact the teachers who were identified by the teacher survey as using the PSI model exclusively in their courses and obtain lesson plans and pacing guides.
B.2	Review lesson plans, pacing guides, and student workbooks to confirm PSI use, as well as for consistency with the weight training curriculum, AKS, and Fitnessgram connections (Table 1). Eliminate teachers from the experimental teacher group who are not consistent.
	D. CONSTRUCT CONTROL GROUP
D.1	Contact the teachers who identified themselves as using the DI model exclusively in their courses and obtain lesson plans and pacing guides.
D.2	Review lesson plans, pacing guides, and student workbooks to confirm non-use of PSI model, as well as for consistency with the weight training curriculum, AKS, and Fitnessgram connections (Table 1). Eliminate teachers from the control teacher group who are not consistent.
	E. DATA COLLECTION
E.1	Collect de-idenfitied rosters with the following data for each student from the PSI teacher group
	a. Pretest Fitnessgram scores: Pacer, pushup, curl up, backsaver sit and reach
	b. Posttest Fitnessgram scores: Pacer, pushup, curl up, backsaver sit and reach
E.2	Collect de-idenfitied rosters with the following data for each student from the DI teacher group
	a. Pretest Fitnessgram scores: Pacer, pushup, curl up, backsaver sit and reach

	b. Posttest Fitnessgram scores: Pacer, pushup, curl up, backsaver sit and reach
E.3	Compare the sample sizes of the experimental and control groups. In the instance of unequal sample sizes, the following method for equalizing sample sizes by random deletion will be followed:
	a. Assign each student within the larger group a three digit number
	b. Use a random number table to construct a sample of equal size to the other group.
F. DATA ANALYSES STEPS	
F.1	Create subgroups based on the research question
F.2	Perform assumption tests for each RQ
	a. Normality - Shapiro-Wilk
	b. Linearity between CV and DV- Scatter plots
	c. homogeneity of variance- Levene's Test for Equality of Error Variances
	d. homogeneity of regression-F test on the interaction of the independent variable with the covariate
	e. reliability of CVs
F.3	Given the one-way ANCOVA, Identify for each research question the significance level, power, effect size, degrees of freedom within groups, degrees of freedom between groups, and the minimum number of participants per group.
	SIGN. LEVEL: Alpha = 0.05
	Power = .8
	Effect Size = d = .5
	MINIMUM PER GROUP- 33 per group
	MINIMUM TOTAL = 33x# of levels of the IV
	Degrees of Freedom Within Groups: N-k-c
	Degrees of Freedom Between Groups: k-1
F.4	Perform ANCOVA and calculate the F ratio for each RQ and hypothesis to confirm or reject the null hypotheses.

APPENDIX B: Dissertation Timeline

Step #	Description	Proposed Month of Completion	Completed
A	**Summer 2012**		
A.1	Secure a Chair		x
A.2	Email prospectus (or Ch. 3), Timeline, and Research Summary to Committee Search		x
A.3	Secure committee members - LUO member		x
A.4	Complete EDUC 980 with a well developed prospectus		x
A.5	Sign up for EDUC 989		x
A.6	Sign up for EDUC 989 - Fall (Pantana)		x
B	**Fall 2012**		
B.1	Start EDUC 989 (Pantana)		x
B.2	Edit prospectus and develop the proposal		x
B.3	Continue to develop Ch.1 - 3 of dissertation		x
B.4	Sign up for EDUC 989 Spring semester		x
C	**Spring 2013- Fall 2014**		
C.1	Secure outside committee member	January-Feb.	x
C.2	Develop proposal and submit to committee for approval to defend proposal		
C.3	Submit paperwork to defend proposal		
C.4	Defend Proposal		
C.5.1	Submit Liberty IRB form & gain approval	Must be done at the same time.	
C.5.2	Submit research request to the Principal using the Local School Research Request form	Deadlines: March 1, April 22, June 24	
C.6	Gain approval from the Principal and fax Local School Research Request form to the district office		
C.8	Contact P.E. department head and receive a list of teachers who teach all-male weight training classes		
C.9	Receive permission from the Principal to administer the teacher survey, administer the teacher survey, and analyze it.		
C.10	Collect lesson plans, pacing guides, and student materials, confirm usages and construct teacher control and experimental groups		
D	**Summer 2015**		
D.1	Collect de-identified rosters from the teacher groups.		
D.2	Create equal sample sizes for each group, and collect de-identified data from the teacher groups.		
D.3	Perform analses on the data		
D.4	Get analyses check by the statistics consultant		
D.5	Continue developing the manuscript		
D.6	Submit manuscript for review		
D.7	Make appropriate changes to the manuscript		
D.8	Sign up for EDUC 989		
E	**Fall 2016**		
E.1	Submit manuscript for review and gain approval to sign up for EDUC 990 D term.		
E.2	Defend Dissertation		
E.3	Make any necessary changes in the manuscript		
E.4	File the proper paperwork to complete the program		

APPENDIX C: IRB Approval Letter

LIBERTY UNIVERSITY.
INSTITUTIONAL REVIEW BOARD

Dear Charles,

The Liberty University Institutional Review Board has reviewed your application in accordance with the Office for Human Research Protections (OHRP) and Food and Drug Administration (FDA) regulations and finds your study to be exempt from further IRB review. This means you may begin your research with the data safeguarding methods mentioned in your approved application and no further IRB oversight is required.
Your study falls under exemption category 46.101(b)(2, 4), which identifies specific situations in which human participants research is exempt from the policy set forth in 45 CFR 46:101(b):

> (2) Research involving the use of educational tests (cognitive, diagnostic, aptitude, achievement), survey procedures, interview procedures or observation of public behavior, unless:
> (i) information obtained is recorded in such a manner that human subjects can be identified, directly or through identifiers linked to the subjects; and (ii) any disclosure of the human subjects' responses outside the research could reasonably place the subjects at risk of criminal or civil liability or be damaging to the subjects' financial standing, employability, or reputation.
> (4) Research involving the collection or study of existing data, documents, records, pathological specimens, or diagnostic specimens, if these sources are publicly available or if the information is recorded by the investigator in such a manner that subjects cannot be identified, directly or through identifiers linked to the subjects.

Please note that this exemption only applies to your current research application, and any changes to your protocol must be reported to the Liberty IRB for verification of continued exemption status. You may report these changes by submitting a change in protocol form or a new application to the IRB and referencing the above IRB Exemption number.
If you have any questions about this exemption or need assistance in determining whether possible changes to your protocol would change your exemption status, please email us at

Sincerely,

LIBERTY UNIVERSITY.

Liberty University | Training Champions for Christ since 1971

1971 UNIVERSITY BLVD. LYNCHBURG, VA. 24515 IRB@LIBERTY.EDU FAX (434) 522-0506 WWW.LIBERTY.EDU

APPENDIX D: Informed Consent Form

The Liberty University Institutional
Review Board has approved
this document for use from
03/30/15 to --
Protocol # 2134.033015

CONSENT FORM

Personalized System of Instruction and Student Performance in High School Weight Training Courses

Charles Allen
Liberty University
School of Education

You are invited to be in a research study that examines the possible impact of the Personalized System of Instruction instructional model on student achievement. You were selected as a possible participant because you taught all-male weight training courses for the 2013-2014 SY at your school. I ask that you read this form and ask any questions you may have before agreeing to be in the study.

Charles Allen, doctoral candidate in the School of Education at Liberty University is conducting this study.

Background Information:

The purpose of this study is to determine the possible impact of the Personalized System of Instruction instructional model on student achievement on the Fitnessgram assessment in high school weight training classes.

Procedures:

If you agree to participate in this study, you will be asked to do the following:

1. Send an email from your personal email address to with the word YES.
2. Collect from your teacher mailbox a folder containing the following items: an informed consent form, a teacher survey, a directions form, and an envelope.
3. Sign the informed consent form and place into the envelope (do not seal yet).
4. Complete a brief teacher survey (5 – 10 minutes) and place it into the envelope.
5. Seal the envelope and place the envelope into the locked cabinet inside the Investigator's teacher mail box.
6. Export your 2013-2014 Fitnessgram pretest and posttest scores for your 2013-2014 classes, including the following Fitnessgram assessments: PACER, 90^0 push-up, curl-up, backsaver sit and reach right and left leg scores for your 2013-2014 weight lifting classes using the **RESEARCH** option (which de-identifies your student data) into an excel spreadsheet (5-10 minutes).
7. Email the exported spreadsheet to (5-10 minutes).
8. Delete the excel spreadsheet from your computer (1-5 minutes).

Risks and Benefits of being in the Study:

No study is without risk. The risks associated with this study are minimal and are no more than a participant would encounter in everyday life.

While there are no direct benefits to participants for participating in this study, the benefits for society are substantial. Research into effective instructional models that can be used in large-sized physical education classes is vital to the continual improvement of our educational system.