Current Practices and Perceptions

of Physical Education

Teacher Evaluation Systems

by

Jason Norris

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved June 2014 by the
Graduate Supervisory Committee:

Audrey Beardsley, Co-Chair
Hans van der Mars, Co-Chair
Pamela Kulinna

ARIZONA STATE UNIVERSITY

August 2014

ABSTRACT

Given the current focus on high-stakes accountability in America's public schools, this study examined teacher evaluation specific to physical education. This study revealed current teacher evaluation practices used in physical education, perceptions of school administrators related to the value of the physical education evaluation process, and the perceptions of the physical education teachers related to the value of the evaluation process. The first phase of this study was an interpretive document analysis study conducted on four separate teacher evaluation systems commonly used within the public school system to evaluate physical education teachers. Those four systems were: Marzanos teacher evaluation model, Danielson framework for teaching (FFT), Rewarding Excellence in Instruction and Leadership (REIL), and Teacher Advancement Program (TAP). A separate evaluation instrument specific to physical education created by the National Association of Sport and Physical Education (NASPE) was used as a comparative evaluation tool. Evidence suggests that two of the four teacher evaluation systems had a high percentage of alignment with the NASPE instrument (TAP 87.5%, FFT 82.5%). The Marzano teacher evaluation model had the least amount of alignment with the NASPE instrument (62.5%). The second phase of this study was a phenomenological approach to understanding administrators' and physical education teachers' perceptions to teacher evaluation specific to physical education. The participants in this study were administrators and physical education teachers from an urban school district. An informal survey and formal semi-structured interviews were used to reveal perceptions of teacher evaluation specific to physical education. Evidence from the administrator's informal survey and formal semi-structured interviews revealed four

i

common themes: (1) "I value PE, but I live in reality" (administrators value physical education, but practice in reality); (2) "good teaching is good teaching"; (3) "I know my limitations, and I want/need help" (relative to teacher evaluation in PE); and (4) where's the training beef? Evidence from the physical education teacher's informal survey and formal semi-structured interviews revealed three common themes: (a) physical education is valued, but not prioritized; (b) teacher evaluation in physical education is "greatly needed, yet not transparent; (c) physical educators are not confident in their evaluator.

DEDICATION

This dissertation is dedicated to my children Jaden and Peyton.

ACKNOWLEDGMENTS

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

The National Association for Sport and Physical Education (NASPE) states that physical education is an integral part of the total education of every child in Kindergarten through Grade 12. The overall goal of a quality physical education program is to produce physically educated persons that adopt healthy and physically active lifestyles (DeJong, Hensley, & Tannehill, 2004). A well-planned and implemented, quality physical education can increase the physical competence, health related fitness, self-responsibility and enjoyment of all students so they can be physically active for a lifetime (NASPE, 2012a). One critical factor in producing these student learner outcomes is having a qualified physical education teacher who can plan and implement such a program.

Teacher evaluation is a standard process used by districts and states aimed at monitoring teacher's performance. In recent years, teacher evaluation in core classroom subjects has been tied more directly to students' performance on standardized achievement tests. This has important implications in terms of having a credible and reliable evaluator to ensure both improvement and accountability of teacher performance (Hill & Herlihy, 2011). With legislation efforts such as Race to the Top (U.S Department of Education, 2009), states and districts have been adopting various teacher evaluation systems in an effort to improve student achievement, decrease the achievement gaps across student subgroups, and increase the rates at which students graduate from high school prepared for college and careers. Moreover, the Race to the Top legislation has caused states to rethink how teachers are evaluated, and to make high stakes decisions such as how teachers are compensated, promoted, granted tenure or dismissed based on

1

their overall effectiveness in the classroom (USDE, 2009). In 2011, the National Council on Teacher Quality (NCTQ) released a report that highlighted the current teacher evaluation and effectiveness policies of all 50 states. Some key findings included the following statistics:

- 24 states and the District of Columbia require annual evaluations for all teachers (p.6)

- 23 states require evidence of student achievement as a preponderant criterion in teacher evaluation (p.6)

- 18 states and the District of Columbia use teacher evaluation results as possible terms for dismissal (p.6)

- 6 of the reported states made post evaluation feedback sessions mandatory (p.18)

While there have been major advances in new state policies surrounding teacher evaluation, the state of the state's report suggest that there are still many shortcomings (NCTQ, 2011). Unlike general education, physical education does not have a tradition of systematic evaluation or formal assessment of student performance (Williams & Rink, 2003). With a current elevated interest in assessment and accountability, assessment has become an area of concern in physical education. The very survival of physical education in the public school system will be determined on the methods used in documenting student and teacher success (Mercier & Doolittle, 2013). To date, there are no known studies that examine teacher evaluation systems specific to physical education. Therefore, this study will examine current teacher evaluation systems and look at measures used in determining quality teaching and effectiveness in physical education. Both administrator

and physical education teachers' perceptions will also be examined in determining the value orientation of the evaluation process.

Background and Problem

  The release of the report "A Nation at Risk (National Commission on Excellence in Education, 1983), first revealed the shortcomings of the United States education system. The commission's report contained several specific areas in which particular attention was to be given. Those areas were 1) assessing the quality of teaching and learning in our Nation's public and private schools, colleges, and universities; 2) comparing American schools and colleges with those of other advanced nations; 3) studying the relationship between college admissions requirements and student achievement in high school; 4) identifying educational programs which result in notable student success in college; 5) assessing the degree to which major social and educational changes in the last quarter century have affected student achievement; and 6) defining problems which must be faced and overcome if we are successfully to pursue the course of excellence in education.

  Two decades later, the federal government passed the No Child Left Behind Act (NCLB, 2002) with the intention of using accountability measures to solidify the U.S.'s chances of eliminating the achievement gap that exist between groups of students within our nation's schools (DeJong, Hensley, & Tannehill, 2004). With the passing and implementation of the NCLB Act, the federal government intensified its focus on accountability directed at student achievement at the school and district levels. With new accountability measures such as state standardized tests, teachers and administrators were put under a microscope and are held responsible for the academic achievement of their

students. The NCLB Acts primary focus is mathematics and language arts that are considered "core subject" areas. Subjects not considered "core subject" areas and with less emphasis on accountability include: science, social studies, music, art and physical education.

Current Status of Physical Education

Since physical education is not regarded a K-12 core subject, state level standardized testing is not a requirement. In 2012, NASPE released a report that revealed the current status of physical education in each state. The purpose of this report was to target the following areas directly related to physical education: 1) high school graduation requirements; 2) exemptions/waivers and substitutions; 3) physical activity; 4) local school wellness policy; 5) standards, curriculum and instruction; 6) class size; 7) student assessment and program accountability; 8) body mass index (BMI) collection; 9) physical education teacher certification/licensure; 10) national board certification in physical education; 11) state physical education coordinator requirements (NASPE, 2012b). Major findings included:

- 74.5% of states mandate physical education in elementary, middle/high, and high school, but most do not require a specific amount of instructional time and more than half allow exemptions, waivers, and/or substitutions (p.7).

- More than half of all states permit school districts or schools to allow students to substitute other activities in lieu of their required physical education credits (p.8).

- Of 49 states responding, 53.1% (26) require physical education grades to be included in a student's grade point average (GPA) (p.8).

- 98.0% of states have adopted their own state standards for physical education, and 76.0% (35 states of 46 survey respondents) require local districts to comply or align with these standards (p.8).

- Of 51 states responding, 50.9% (26) require some form of student assessment in physical education (p.8).

- 68.6% of 51 states reporting require those who desire to teach physical education to pass a certificate/licensure exam before they can teach physical education (p.8).

- 24 states (47%) require professional development for physical education teachers on physical education topics, comparable to other curricular areas (p.9).

- 19.6% (10 of 51 states) provide any funding for professional development that is specifically for physical education teachers (p.9).

Although these statistics show slight improvements from a previous report released in 2010 (NASPE, 2010), physical education remains a marginalized subject with little to no accountability for the delivery of quality and sufficient quantity of it in schools.

Health and Physical Activity Status among Children and Youth

Currently, the U.S. suffers from what the Centers for Disease Control and Prevention (CDC) has described as an obesity epidemic (CDC, 2011a). The CDC has reported that childhood obesity affects 17% of all children and adolescents in the U.S. That is triple the rate from one generation ago (CDC, 2012). Childhood diabetes is now also on the rise (Levi, Segal, & Juliano, 2013). Over the past two decades, there has been a rise in the detection of type 2 diabetes (also known as Adult onset diabetes) among U.S. children and adolescents (CDC, 2012). The U.S. Surgeon General has recommended at

least 60 minutes a day and 150 minutes a week of moderate to vigorous physical activity (MVPA) (USDHHS, 2008).

Based on the most recent Youth Risk Behavior Surveillance System(YRBSS) results, 14% of high school students do not participate in 60 minutes or more of daily physical activity, and 32% report playing video games or watching television for three or more hours a day (CDC, 2011b). At the same time nearly one-third of children and teens, more than 23 million kids, are overweight or obese, and physical inactivity is a leading contributor to the epidemic (Trost, 2009). Students' activity levels are declining as obesity related diseases increase in prevalence. Given the amount of time children spend at school, that time is crucial in receiving opportunities to be physically active.

The Importance of Physical Activity and Physical Education's Role

Physical education is one of only a few interventions for which there is sufficient evidence to support the outcomes of increasing the moderate to vigorous physical activity levels in children and youth (CDC, 2001). Thus, it is more important than ever that physical educators deliver quality and standards-based physical education to their students. Quality physical education programs offer students learning opportunities focused on health related fitness, physical competence, cognitive understandings, and positive attitudes about physical activity (PA)(Sallis & McKenzie, 1991; Sallis et al., 2012).

Teacher Evaluation and Accountability

With no accountability measures or standardized testing tied to assessment of quality teaching in physical education, formal teacher evaluation systems serve as the primary source of determining teacher performance. Teacher evaluations serve two primary

6

purposes: improvement and accountability (Koppich, 2008). Boyd (1989) states that effective teacher evaluation systems should have the following procedures and standards:

1. Relate to important teaching skills

2. Be as objective as possible

3. Be clearly communicated to the teacher before the evaluation begins and be reviewed after the evaluation is over

4. Be linked to the teacher's professional development

Good evaluation is a continuation of good professional development (Koppich, 2008). Effective professional development is standards and content based, to a large extent teacher-driven, closely aligned to what teachers do in their schools and classrooms, and part of each teacher's workday (Koppich, 2008). Sadly enough, current evaluation practices often don't align with curricular standards, and professional development efforts do not result in focused instructional support (Heneman, Milanowski, Kimball, & Odden, 2006). With a lack of instructional support and non-alignment of content curricular standards, opportunity for teacher improvement by way of teacher evaluation may be missing in most cases.

Teacher evaluation systems not only can contribute to the professionalization of teaching, but will also serve to invest educators with greater information, confidence, and ability to improve their instructional practices, on top of helping students achieve their fullest potential (Goldrick, 2002). In making sure that students are receiving the maximum opportunity to learn, districts use teacher evaluation to ensure quality teaching is present in the classroom. This accountability or quality assurance is to ensure each classroom is equipped with a competent teacher (Danielson, 1996). Teaching is a

complex activity that needs much more than brief observations of a teacher. Effective

evaluation systems should recognize, cultivate and develop good teaching (Danielson,

1996).

Teacher evaluation has the potential to improve instructional effectiveness and

student learning by enabling teachers to receive high-quality guidance and feedback, thus

improving their instruction (Donaldson, 2009). With so many high stakes decisions being

made in regard to teacher and student performance in the classroom, this study looks

mainly at professional growth and instructional practice within the physical education

teacher evaluation process.

Current Trends in Teacher Evaluation

Value Added Assessment Systems. Since the introduction of the No Child Left

Behind Act in 2002, and President Obamas' Race to the Top initiative in 2009, value

added measures of accountability have been adopted to determine the value a teacher

adds to student learning. These measures are referred to as value added models (VAMs).

An example of a VAM and the most commonly used model today is the SAS Education

Value Added Assessment System (SAS EVAAS). This particular assessment system is

used to collect test score data and measure learning trajectory from the time the student

enters the teacher's classroom to the time they leave (Amrein-Beardsley & Collins,

2012). Rothstein et al (2010) noted that while value-added models (VAMs) contribute to

stronger analyses of school progress, program influences, and increased validity of

evaluations, these methods alone are not reliable and valid indicators of teacher

effectiveness. Herlihy (2012) examined state and local efforts to investigate validity and

reliability of scores from teacher evaluation systems. Few states seemed to be considering

the negative unintended consequences of systems that may generally be perceived as being arbitrary. Furthermore, administrators from many states understand the importance of implementing better systems, but lack knowledge of implications for validity and reliability of scores produced by their systems (Herlihy, 2012). Moreover, Berliner (2014) pointed out the many exogenous variables associated with VAM's, indicating that the major problem for value added approaches is assessing teachers based upon student outcomes, when countless variables beyond the classroom affect achievement inside the classroom.

With student learning having become a major concern within our schools, various teacher evaluation systems have been created aimed at presenting more valid judgments towards determining teacher effectiveness. These systems contain standards and detailed rating scales, which provide guidance to evaluators in determining teacher performance (Kimball & Milanowski, 2009). The Danielson "framework for teaching" is a widely used teacher evaluation system (Danielson, 1996). Moreover, the framework for teaching system contains 22 components within four domains of teaching practice: planning and preparation, classroom environment, instruction and professional responsibilities. There are 66 elements that list various aspects of performance on the components and domains. There is a four level rubric that provides a range of proficiency measured by observed performance descriptions. These levels range from unsatisfactory to distinguished teaching practice. This system is generic in that it purports to apply to all grade levels and subject areas (Kimball & Milanowski, 2009). Danielson's system has been validated and studies have shown improved student learning (Danielson, 2012). Other popular used systems include: (a) the teacher advancement program (TAP) (National Institute for

Excellence in Teaching, 2013); (b) the Marzano teacher evaluation model (Marzano Research Laboratory, 2013); and (c) rewarding excellence in instruction and leadership (REIL) (Maricopa County Education Service Agency, 2013). The use of the aforementioned instruments vary from state to state and district to district.

Heneman and colleagues (2006) stated that for an evaluation system to be credible and useful, the following five procedures must be met: a) Establishing accepted, evidence-based teaching standards; b) using a valid instrument; c) thoroughly training and recalibrating raters; d) employing multiple evaluators; and e) establishing a process for providing feedback and targeting support. They also claim that the evaluation system itself must promote transparency so that teachers can easily understand expectations, and also serve as a way to increase communication between evaluators and evaluates. Teachers need to feel benefits from the system rather than only judged (Heneman et al., 2006).

The Pitfalls of Teacher Evaluation Systems

With so much emphasis on improving standardized testing scores and year-to-year student growth, teachers are defined and held accountable for these results (Hinchey, 2010). Using a single measurement in making high stakes decisions such as termination, pay, and tenure, has become a major issue within our public school system. There is evidence that value added measures in teacher evaluation systems may be invalid and unreliable sources in determining teacher quality and student achievement (Amrein-Beardsley & Collins, 2012; Hill & Herlihy, 2011). Furthermore, tying sanctions to test scores and value added measures can potentially yield consequences such as discouragement of teachers wanting to work in underachieving schools, along with

10

demoralization of teaching in general (Rothstein et al., 2010). Marion and Buckley

(2011) pointed out that many Race to the Top applications included promises that States

would use other forms of data to demonstrate student achievement in non-tested subjects.

This poses a potential problem for physical educators based upon lack of national

standardized assessment measures to demonstrate student learning, along with the unfair

assumption that physical education teachers are sharing responsibility of student

achievement on a school-wide level.

Statement of the Problem

A major problem with physical education teacher evaluation is linking evaluation

outcomes with teaching practice and now also with student learning data. Moreover,

current evaluation practices do not align with curricular standards and professional

development efforts and do not result in targeted instructional support (Heneman et al.,

2006).

A second problem is that teacher evaluations are not always conducted by a

person qualified and trained to do so, which may lead to subjective and bias scoring

outcomes (Brandt, Mathers, Oliva, Brown-Sims, & Hess, 2007; Little, 2009). This lack

of credibility becomes an issue when the evaluator is unfamiliar with observation

techniques, does not know criteria to follow, and does not use effective methods to share

feedback from the evaluation process (Brandt et al., 2007).

Finally, there is also a problem of the likely lack of content and pedagogical

knowledge specific to the content (in this case physical education) that makes teacher

evaluation especially difficult for the school administrator (Donaldson, 2009). Not only

does this increase the risk of invalidating the administrators' observations, judgments and

final ratings, it could also lead to missed professional growth opportunities, lowered student learner outcomes, along with less value placed on physical education teacher evaluation as a whole, from both teacher and evaluator.

Human Resource Theory as a Theoretical Framework

The school is an organization that seeks the successful growth and achievement of the students that are being educated within its walls. The teachers are the primary focus of responsibility for students' learning outcomes. Furthermore, the growth and development of teachers is a potential facilitator of student learning. The human resource theory focuses on the abilities of an organization to succeed at the growth and development of its employees in progressing towards common goals (Argyris, 1970). Human resource theorist assumes the following:

- Organizations exist to serve human needs.

- Organizations and people need each other.

- When the fit between the individual and the organization is poor, one or both will suffer; individuals will be exploited, or will seek to exploit the organization, or both.

- A good fit between individual and organization benefits both: human beings find meaningful and satisfying work, and organizations get the human talent and energy that they need (Bolman & Deal, 2003).

Human resource theory suggests that what is good for the organization, must be good for the individuals within that organization. Therefore, a healthy relationship between the school and the teacher may lead to achievement of desired goals. Moreover, when there is

a balance of individual needs with institutional expectations, there will be a fostering of productive work environments (March & Simon, 1993).

Purpose of the Study

Given the current focus on high–stakes accountability in America's public schools, and the status of physical education within that context, the purpose of this study is to: a) gain a better understanding of current teacher evaluation practices used in physical education; b) reveal perceptions of school administrators related to the value of the evaluation process, specific to physical education; and c) reveal perceptions of physical education teachers related to the value of the evaluation process.

Research Questions

The following are three specific research questions for this study:

1. What are the current teacher evaluation documents school administrators' use when conducting formal evaluation of physical education teachers?

2. How do the administrators view the evaluation process specific to physical education?

3. How do the physical education teachers view the evaluation process?

CHAPTER 2

LITERATURE REVIEW

There is evidence that quality instruction is connected to positive academic outcomes in student learning (Darling-Hammond, 2000). Furthermore, without high quality evaluation systems, we cannot know if we have high quality teachers (Stronge & Tucker, 2003). With the current push for more accountability within our schools and an emphasis on teacher effectiveness, teacher evaluation systems have become a major focus within educational reform (Danielson, 1996). While a majority of educators believe that teacher evaluation is geared towards improving instruction and teaching practices, legislations and other stakeholders call for accountability and minimum levels of performance (Danielson & McGreal, 2000).

While current trends in teacher evaluation are more focused on core subject areas (NCLB, 2002), other subjects such as physical education, art and music have received minimal interest and allocation of time towards accountability measures. Specific to physical education, there are a limited number of states in the US that have a formal assessment and evaluation system that hold school districts, schools and teachers accountable for students meeting state and national standards. According to the National Association of Sport and Physical Education (NASPE), 50 U.S. states (98%) have developed content standards that reflect those set by NASPE or locally developed state initiatives(NASPE, 2012). However, only 35 states (76%) require local districts to comply or align with these standards and only 26 states (50.9%) mandate some form of student assessment in physical education.

With minimal studies examining teacher evaluation and accountability in non-core subject areas, this review will examine current teacher evaluation systems along with emphasis on value and perceptions of those most revealed with these systems.

Current Teacher Evaluation Systems

While many teacher evaluation systems are being used throughout the country, this section will introduce four systems commonly used within the U.S. These four systems are: a) teacher advancement program (TAP); b) rewarding excellence in instruction and leadership (REIL); c) the framework for teaching (FFT) (Danielson, 1996; Danielson, 2007); and d) the Marzano teacher evaluation model (Marzano, 2003).

The teacher advancement program (TAP) was created by educational reformer Lowell Milken of the Milken Family Foundation. The TAP program is constructed in such that it works in improving the recruitment and retention of talented teachers by restructuring the evaluation and rewards system within the school (Little, 2009). The TAP system links accountability with compensation and has a comprehensive approach, focusing on 4 key elements: a) Multiple career paths, b) ongoing applied professional growth, c) instructionally focused accountability, and d) performance-based compensation. Once TAP is adopted within a school, there is a one-year grace period in which administrators and teachers learn the program. Moreover, after one year's time, the evaluation system begins, and monetary consequences are put into place. This system targets improving teacher practices and increasing student learning both individually and collectively as a school. In two recent studies, TAP was compared with control schools in use of other evaluation systems. TAP schools teachers consistently outperformed the others in both student achievement and proficiency levels (Schacter, & Thum, 2004;

Solomon, White, Cohen, & Woo, 2007). In two separate studies, Principals and teachers received surveys, both showing positive results when asked about their satisfaction levels of the TAP program (Agam, Reifsneider, & Wardell, 2006; Agam & Wardell, 2007).

The rewarding excellence in instruction and leadership (REIL) evaluation system is similar to that of TAP, in that it ties teacher compensation to evaluation outcomes and student level of success. REIL is a five-year initiative that was funded by the Teacher Incentive Fund through the US department of education, and is specific to Maricopa County in the Western United States (MCESA, 2013). REIL targets five critical elements of teacher evaluation: a) rigorous, fair and transparent educator evaluations; b) targeted professional learning; c) tools for measuring student success; d) establishment of multiple career pathways; and e) sustainable, differential, performance-based compensation. Moreover, REIL was designed using a sound body of research that targeted effective measures of teaching, student achievement, and administrative leadership strategies (Bill & Melinda Gates Foundation, 2013; Hussey & Khandaker, 2012; Chait & Miller, 2010). The REIL system calls for observation cycles by a trained and certified evaluator five times over the course of the school year. Teachers have the opportunity to meet with their evaluator prior to the formal observation in a "pre-conference" at which lesson plans and objectives, along with expected outcomes are discussed. Furthermore, teachers will then be given the opportunity to meet in a "post-conference" meeting at which reflections of the evaluation from both teacher and evaluator will occur. Supporters of REIL believe that once funding has ceased upon the five-year window, sustainability will be possible due to the positive impact on instruction, student achievement and professional growth of administrators (MCESA, 2013).

Charlotte Danielson (1996) created the framework for teaching (FFT) as a way to improve teacher instruction. The FFT is a widely used system and recognized nationwide (Little, 2009). Danielson's model has four domains that target different areas of teaching performance. Those four domains are: a) planning and preparation; b) classroom environment; c) instruction; and d) professional responsibility. Within each of the four domains are 22 specific performance components, which then have 76 specific task elements that are measured on four levels of performance; unsatisfactory, basic, proficient, and distinguished. Furthermore, the FFT is designed to cover all instructional content areas at grade levels k-12 (Danielson, 2007). FFT can be used for both summative and formative assessments, but is also used for other purposes that promote professional growth. Little (2009) stated "FFT serves as a useful framework with which to link together improvement, evaluation, and other human capital development activities". Henneman, Milanowski, Kimball, & Odden (2006) consolidated research studies conducted over four sites. The outcomes of these studies all showed positive correlations between FFT scores and student achievement, which were measured across value added gains on standardized test. Out of the four studies, significant gains were measured in two particular sites. The two sites differed from the others in that rigorous high-quality training was mandatory prior to evaluations, also that sites included multiple evaluators (Henneman et al., 2006).

The Marzano evaluation model is comprised based upon a number of previous and related works (Marzano, 2007; Marzano, 2006; Marzano, 2003; Marzano, Frontier, & Livingston, 2011; Marzano, Pickering, & Marzano, 2003; Marzano, Pickering, & Pollock, 2001). The Marzano model includes four domains that target different areas of

teaching. Those four domains are: a) classroom strategies and behaviors; b) preparing and planning; c) reflecting on teaching; and d) collegiality and professionalism. Moreover, the domains include 60 elements that target various behaviors specific to teaching (Marzano Research Labortory, 2011). The Marzano system lays out a road map for evaluators and teachers to follow over the course of the school year. Thus, specific meetings, self-reflections, observations, goal setting and feedback sessions are required in following the Marzano system (Marzano Research Laboratory, 2011). Marzano and colleagues (2012) developed a summary discussing the results of four separate studies all aimed at examining the role of the Marzano model of teacher evaluation in student achievement. In all four studies, positive correlations surfaced between teachers' use of the Marzano model and students' learning and achievement (Haystead, 2010; Marzano Research Laboratory, 2010; Marzano & Haystead, 2010; Marzano & Haystead, 2011).

All four of the aforementioned evaluation systems target important areas specific to teaching behaviors present within the classroom. The specific areas are determined in a universal and generic format. The TAP, FFT, and Marzano teacher evaluation systems have substantial amounts of research findings revealing positive student achievement within core subject classrooms when faithfully implemented within their schools. Moreover, there were no current studies identified that examined the effects on teacher and student achievement within the physical education classroom, using a standards based teacher evaluation system.

While there tend to be studies that support the aforementioned teacher evaluation models and their effectiveness in the classrooms, there is virtually no empirical evidence

of rigorous peer reviewed studies to support improved teacher performance and student outcomes (Peterson, 2000).

Value-Added Models of Teacher Evaluation (VAM'S)

Value-added models (VAMs) are widely used to estimate student's growth from year to year. The No Child Left Behind Act (2002) laid the groundwork for teachers to be held more accountable for student outcomes (Braun, Chudowsky, & Koenig, 2010). Rothstein et al (2010) noted that while value-added models (VAMs) contribute to stronger analyses of school progress, program influences, and increased validity of evaluations, these methods alone are not reliable and valid indicators of teacher effectiveness. Herlihy (2012) examined state and local efforts to investigate validity and reliability of scores from teacher evaluation systems. Few states seemed to be considering the negative unintended consequences of systems that may generally be perceived as being arbitrary. Furthermore, administrators from many states understand the importance of implementing better systems, but lack knowledge of implications for validity and reliability of scores produced by their systems (Herlihy, 2012). Finally, there is evidence that points out the many exogenous variables associated with VAM's, indicating that the major problem for value added approaches is assessing teachers based upon student outcomes, when countless variables beyond the classroom affect achievement inside the classroom (Amrein-Beardsley, & Collins, 2012; Berliner, 2014; Hill, Kapitula, & Umland, 2011).

Administrator's Role in Teacher Evaluation

Researchers have examined the administrator's role in teacher evaluation since the 1970's (Blase & Kirby, 2000). Principals who support and initiate a democratic

working environment, in which teachers have increased participation in decision making and leadership roles, generally gain the loyalty, trust and respect of their faculty (Allen, Glickman, & Hensley, 1998; Zimmerman & Deckert-Pelton, 2003). Furthermore, principals who support collaboration, open communication, and focus on promotion of professional development amongst their teachers, have a more positive impact on their pedagogical skills (Zimmerman & Deckert-Pelton, 2003; Blase & Blase, 1998; Conger & Kanungo, 1994; Sheppard, 1996). Evidence indicates that teachers respond better to administrators when they use human relation skills, rather than using the power of authority (Treslan & Ryan, 1986). These improved relations support positive working environments, as teachers feel a sense of belonging within an organization (March & Simon, 1993).

The administrators' attitude towards teacher evaluation can affect the evaluation process in determining accuracy and validity of teaching practices (Tziner, Murphy, & Cleveland, 2001). The administrators' skill level at identifying teaching in unfamiliar content areas is very important. The more skilled the evaluator, the more accurate the scores on determining effective teacher performance on the dimensions defined by the evaluation system (Kimball & Milanowski, 2009). Research indicates that evaluator training in all areas of content is necessary in determining teacher effectiveness and overall evaluator accuracy (Sartain, Stoelinga, & Krone, 2010).

Kersten and Israel (2005) conducted a study examining principals' perceptions of the benefits of teacher evaluation. While a majority of principals devote a considerable portion of their time implementing the district required teacher evaluation system, they generally perceive that the process has a limited direct impact on improving an

individual's teaching and subsequent student learning (Kersten & Israel, 2005). The

principal's pointed out several benefits of teacher evaluation including: 1) goal setting;

2) enhanced supervision; 3) enhanced communication; and 4) comprehensive process

(Kersten & Israel, 2005). While benefits were identified, impediments to highly effective

teacher evaluations were pointed out as well. Some areas revealed were: a) time; b)

unions; and c) school culture. These areas of impediments may lead to non-thorough,

dishonest and an erosion of confidence, in the evaluation process (Kersten & Israel,

2005). Other barriers perceived by principals were teacher tenure and restrictive rules of

evaluation (Painter, 2001).

   With many high stakes decisions based upon the outcomes of teacher assessments

and evaluation, reliability has become clearly problematic across principals (Kimball &

Milanowski, 2009). Moreover, Kimball and Milanowski (2009) stated that "providing

evaluators with relatively detailed rubrics or rating scales describing generic teaching

behaviors thought to promote student learning, coupled with initial training in applying

them, is not enough to ensure that all evaluations ratings will be positively related to

student achievement"(p.65). The lack of reliability along with subjective outcomes has

often times been blamed on the principal. Furthermore, Calabrese, Sherwood, Fast, &

Womack (2004) found that principals themselves have ill fillings towards the evaluation

process and have felt as if they were the victims. Moreover, principals in the Calabrese et

al study also felt as if they had no control over what evaluation system were being used

and they were forced to participate regardless of training and familiarity with the

instrument (Calabrese et al., 2004). Principals have also shown concerns with lacking in

the necessary subject-area knowledge for all disciplines. This has shown to be a negative

influence on the strength of evaluation outcomes and the ability to provide instructional feedback (Painter, 2001).

Evaluator training and credibility remain an ongoing issue in teacher evaluation. Most district policies provide little guidance on consistency of evaluations, criteria to follow, and how to use and share feedback from the process (Brandt, Mathers, Oliva, Brown-Sims, & Hess, 2007). In a study conducted to determine principals' perceptions of rating scales as part of the evaluation process, they felt that subjectivity was exacerbated by the lack of a definitive definition for each category (Calabrese et al., 2004).

Teachers Perceptions of the Evaluation Process

Teacher evaluation can be an effective tool in continual pedagogical improvement (Atkins, 1996). For teacher evaluation to demonstrate successful outcomes, there needs to be a level of mutual trust and understanding between the teacher and evaluator. This communication and trust needs to be reciprocal in achieving not only improved pedagogical skills, but an overall positive attitude towards the teacher evaluation in general (Davis, 1988; Valentine, 1992). Zimmerman (2003) conducted a study to gain understanding of teachers' perceptions of evaluation. The teacher's pointed out that communication is one of the key elements inherent in the teacher evaluation process. Forty nine percent of teachers surveyed stated that "feedback, negative or positive, would be welcome, and it must be constructive to be effective". Few teachers felt a relationship of positive communication between themselves and their administrators. They felt that principals were just "filling in the squares, and having them sign on the dotted line" (Zimmerman, 2003). Evidence also revealed a lack of connection between the teacher and the evaluation process itself. Many teachers felt the evaluation process was not

tailored to the individual, but more of a generic approach. Teachers wanted tailored

feedback geared towards their particular needs for improvement. Teachers in Zimmerman

(2003) study stressed the importance of verbal feedback rather than written feedback in

order to have input before, rather than after the fact (Zimmerman, 2003).

Many teachers question the validity of the teacher evaluation instruments, the

subjectivity, and have an overall lack of confidence in the evaluator (Calabrese et al.,

2004, Donaldson, 2009; Little, 2009; Prince et al.2008; Zimmerman & Deckert-Pelton,

2003; Zimmerman, 2003). Kennedy (2008) noted that there are many qualities and

practices that are assessable in teachers, but what is lacking are strategies for organizing

assessments into a coherent system. Moreover, the author claimed that the challenges lie

in the assessment, both what and how to assess, along with organization into a

comprehensive, multifaceted system (Kennedy, 2008). In a recent study measuring

teacher perceptions of the evaluation process, some teachers stated that principals

purposely did not assign *exceeds expectations* as part of a perceived district policy

(Calabrese et al., 2004). This notion that "no one" receives *exceeds expectations* led

teachers to believe their scores were subjective, questioning not the principal, but the

instrument itself. Ovando (2001) revealed similar findings of subjectivity and scoring

bias. Teachers were disappointed in the limited use of *distinguished,* the highest rating

(Ovando, 2001). In other studies, the evaluator herself was held in question as to whether

or not she was qualified to evaluate content based pedagogy (Brandt et al., 2007;

Halverson et al., 2004; Little, 2009; Prince et al., 2008) In a study measuring teacher

perceptions of the administrator's role in evaluation, teachers felt their principals were

not adequately qualified to evaluate the subject area. Moreover, evaluators who lacked

instructional skills were not perceived as having the ability to evaluate instructional

content decisions or pedagogical content knowledge (Halverson, Kelley, & Kimball,

2004).

The review of the literature surrounding teacher evaluation has revealed the need

for continued examination of this topic in many different areas. The literature review has

revealed the importance of peer reviewed research needed in examining the impact

commonly used teacher evaluation systems have on teacher effectiveness and student

outcomes. The literature review has revealed evidence that there is disconnect between

the teacher and the evaluator in communication, expectations during an evaluation, and

post evaluation feedback. Other evidence from the literature pointed out the lack of

confidence teachers have in their evaluators and also the lack of trust the teacher's had in

the evaluation instrument itself. Finally, the literature review has revealed that there are

no known studies that examine teacher evaluation specific to physical education. With

the current push for the use of value added models and more evidence based learning

outcomes, the need to examine teacher evaluation in physical education is greatly needed.

CHAPTER 3

A DOCUMENT ANALYSIS OF TEACHER EVALUATION SYSTEMS SPECIFIC TO

PHYSICAL EDUCATION

Teacher evaluation serve two distinct purposes: (1) personnel decisions (e.g. tenure, termination); and (2) improvement of practice (i.e. professional development) (Scriven, 1981). With the current push for more accountability of teachers in the classroom (United States Department of Education, 2002, 2009), showing evidence of student outcomes and teacher effectiveness has become a priority. Thus, the use of various teacher evaluation systems have been adopted by districts and may lead to high stakes decisions such as job termination. Furthermore, it is very important that high quality teacher evaluation systems are used to provide teachers with the tools they need to continuously tailor instruction, enhance practice, and advance student learning (National Education Association, 2014). The literature addressing the purpose and importance of high quality teacher evaluation systems served as a conceptual framework for this document analysis (Boyd, 1989; Henneman et al., 2006; Koppich, 2008; Scriven, 1981).

Four commonly used teacher evaluation systems were examined in this study. They included: (a) teacher advancement program (TAP) (National Institute for Excellence in Teaching, 2013); (b) rewarding excellence in instruction and leadership (REIL) (Maricopa County Education Service Agency, 2013); (c) the framework for teaching (FFT) (Danielson, 1996, 2007); and (d) the Marzano teacher evaluation model (Marzano Research Laboratory, 2013; Marzano, 2003). All four of these systems target important areas related to quality teaching, and all four emphasize student achievement.

25

Teacher Advancement Program (TAP)

The TAP program was created to assist with recruitment and retention of talented teachers in restructuring the evaluation and rewards system within school (Little, 2009). The system links accountability with compensation by focusing on the following four key elements: (a) multiple career paths; (b) ongoing applied professional growth; (c) instructionally focused accountability; and (d) performance-based compensation. School districts adopt this system, and after one year time, monetary consequences are put into place (National Institute for Excellence in Teaching, 2013). Many studies have been conducted that examine TAP and the effects it has teachers professional growth when implemented in schools. Moreover, studies have indicated that teachers in TAP schools consistently outperform teachers in control schools across many states (Schacter, & Thum, 2004; Solomon et al., 2007). In an evaluation of the TAP program reported in 2010, Glazerman and Seifullah stated that much of the existing evidence on the effects of TAP had been conducted by the programs developers. Furthermore, because of the convenience of self-selected samples, there is great possibility of program bias calling for a need for more evidence (Glazerman & Seifullah, 2010).

Rewarding Excellence in Instruction and Leadership (REIL)

The rewarding excellence in instruction and leadership (REIL) teacher evaluation system is similar to that of TAP in that it ties teacher pay to evaluation outcomes and student success. REIL is a five-year initiative that was funded by the Teacher Incentive Fund through the US department of education, and is specific one particular county in the Western United States (MCESA, 2013). REIL targets five critical elements: (a) rigorous, fair and transparent educator evaluations; (b) targeted professional learning; (c) tools for

measuring student success; (d) establishment of multiple career pathways; and (e) sustainable, differential, performance-based compensation. The REIL teacher evaluation system was developed using a sound body of research from other performance based systems created by Teacher Incentive Fund grantees (Chait & Miller, 2010; Bill and Melinda Gates Foundation, 2013; Hussey & Khandaker, 2012;). Because the REIL is specific to certain districts of one County in the Western U.S. and not commonly used outside of the state, there are no known studies that examine its effectiveness to date.

Charlotte Danielson Framework for Teaching (FFT)

Charlotte Danielson's FFT was created in 1996 as a way to assist in the improvement of teacher instruction (Danielson, 1996). The system has four domains that target different areas of teaching: (a) planning and preparation; (b) classroom environment; (c) instruction; and (d) professional responsibility. Within each of these four domains are 22 specific performance components, which also include 76 specific task elements that are measured using a rating scale: (a) unsatisfactory; (b) basic; (c) proficient; and (d) distinguished. Over the years, many studies have been conducted that show positive correlations with teachers evaluation scores and student achievement (Bill and Malinda Gates Foundation, 2013; Kane, Taylor, Tyler, Wooten, 2010; Sartain, Stoelinga, & Brown, 2009). In a study spanning over four states, researchers found positive correlations between FFT scores and student achievement (Henneman, Milanowski, Kimball, & Odden, 2006). Although the Danielson FFT system is supported by many studies validating its effectiveness, the majority of them are not published in peer reviewed journals.

Marzano Teacher Evaluation Model

The Marzano teacher evaluation model is comprised of items based upon previous works associated with teaching areas and behaviors (Marzano Research Laboratory, 2013). Marzanos' model includes four domains that target different areas of teaching performance. They include: (a) classroom strategies and behaviors; (b) preparing and planning; (c) reflecting on teaching; and (d) collegiality and professionalism. Moreover, the domains include 60 elements that target various behaviors related to teaching.

The Marzano system maps out a route for both administrators and teachers to follow over the course of the school year. Thus, specific meetings, self-reflections, observations, goal setting and feedback sessions are required to stay on track. According to the Marzano system's website (http://www.marzanoevaluation.com), the system is research based and has many studies that show positive correlations between use of the model with teachers and student achievement (Haystead, 2010; Marzano Research Laboratory, 2010; Marzano & Haystead, 2010; Marzano & Haystead, 2011). Most of the studies examining the Marzano system have been conducted and published by Marzanos Research group. Thus, further independent and peer-reviewed studies may be needed to determine the effectiveness of this model.

Teacher Evaluation

Darling-Hammond (2000) determined that quality teachers make a great impact on student learning within the classroom. However, determining quality teaching and identifying teacher pedagogical skills may be a difficult endeavor based upon the extreme variability of student aptitude and prior instruction (Zimmerman & Deckert-Pelton, 2003). Furthermore, in order to effectively identify quality teaching, knowing what to

measure and how to measure it is very important (Kennedy, 2008). The most common method of teacher observation/evaluation is done by the school principal/administrator. Typically, a district's evaluation document provides little guidance on what to observe across subject areas and how to analyze the outcomes and provide feedback to the teachers (Brandt, Mathers, Oliva, Brown-Sims, & Hess, 2007). Historically, teacher assessment has been used to weed out underperforming teachers (Halverson, Kelley, & Kimball, 2004). Thus, current evaluation practices lack consistency in measuring teaching effectiveness, which has led to a system in which 90 percent of teachers are labeled as top performers, with only a small number labeled unsatisfactory (Little, 2009). Moreover, with such a high number of teachers deemed as top performers, those teachers lacking in pedagogical content knowledge and teaching effectiveness may slip through the cracks (National Education Association, 2014).

Henneman, Milanowski, Kimball, and Odden (2006) suggested that credible evaluation systems need to have five crucial elements: (a) evidence-based teaching standards; (b) valid instrumentation; (c) thorough training and recalibrating of raters; (d) multiple evaluators; and (e) established process for providing feedback and targeting support. Furthermore, teacher evaluation systems need to facilitate increased communication and be transparent so that teachers know exactly what is expected of them (Henneman et al., 2006). Current evaluation systems are structured so that scoring rubrics target teaching components that are generic. That is, teaching skills that are deemed to apply to "all" subject areas (Danielson, 1996; Danielson, 2007; Little, 2009; Marzano Research Laboratory, 2011; MCESA, 2013). Furthermore, generic instrumentations may be invalid in measuring teacher quality and performance in

physical education, for the simple fact that dynamics differ from general classroom to physical activity area settings.

Value added models (VAM's) are assessments used to link academic growth with a particular teacher. These models are associated with tested subjects and used to hold teachers accountable by the use of complex formulas predicting the amount of academic growth of a student in a given year. The validity and reliability of scores produced by VAM's have been questioned due to the lack of consideration of multiple variables associated with students' academic growth both in and outside of the classroom (Amrein-Beardsley & Collins, 2012; Berliner, 2014; Hill, Kapitula, & Umland, 2011).

Teacher evaluation in physical education. Unless evaluators are trained and fluent in physical education content and reliability issues, bias-scoring outcomes may be an issue (Brandt et al., 2007). Bias scoring occurs when evaluators do not have adequate training and are unaware of the characteristics and behaviors that the evaluation is designed to measure (Olivia, Mathers, & Laine, 2009).

Physical education is considered a non-core subject and is not subject to standardized testing procedures (NCLB, 2001). Moreover, current trends in value added measures of student achievement are either nonexistent in physical education or used based upon other school wide measures of achievement (Prince et al., 2008).

With little to no accountability measures enforced in physical education, formal teacher evaluations serve as the sole indicator of teacher quality and effectiveness (NASPE, 2012). Therefore, the purpose of this document analysis study was to examine current teacher evaluation systems, understand current practices, and determine whether

the instrumentation used is a valid measure of teaching quality as reflected in teacher behavior and effectiveness specific to physical education.

Methods

Participants and Setting

An interpretive document analysis study was conducted on four separate teacher evaluation systems collected from 10 school districts from the Western United States. The districts were located within the largest county in their state, and consisted of rural, suburban, and urban school districts. The county was chosen for this study due to its diversity in demographics, along with the convenience in location for the research team.

Data Collection

Recruitment. Recruitment letters were sent to the superintendent's office of 56 school districts within the Western United States (See Appendix A) requesting that any and all documents associated with their teacher evaluation systems for physical education teachers be shared with the research team. To increase response rates, the researcher called each district office and spoke directly with a curriculum coordinator or representative responsible for teacher evaluation Of the 56 school districts, 10 districts responded and supplied their evaluation system information. From these documents, four common evaluation systems were identified and served as the data sources for this study.

Comparative Evaluation Tool

The National Association for Sport and Physical Education (NASPE, 2007) created a teacher evaluation tool to identify the knowledge, skills, and behaviors needed to provide sound instruction in the k-12 physical education classroom. The tool is used as a resource for evaluating teacher behaviors and effectiveness in physical education

settings. The NASPE tool consists of five domains: (1) Instruction; (2) Evidence of Student Learning; (3) Management/Organization; (4) Learning Climate; and (5) Professionalism. Within each domain, there are multiple elements that reflect different teaching behaviors. In total, there are 67 elements within the 5 domains.

The NASPE physical education teacher evaluation instrument was used in this study as a basis for resource to determine whether or not the targeted knowledge, skills and behaviors within physical education settings were present within the four teacher evaluation systems being used in the state. Domain five (Professionalism) was not used in this study due to non-observable teaching behaviors targeted in its elements.

Data Analysis

Two rounds of document analyses were used to ensure an accurate depiction of the data.

Key items from NASPE instrument. In the first round of analysis, the researcher and another independent reviewer examined and discussed each of the domains and elements within the NASPE evaluation tool. Based on overlapping of elements and non-observable teaching behaviors (e.g. class planning and preparation) the researchers narrowed the list of 67 key items (elements) down to 55. The researchers discussed the 55 key items from the NASPE tool and determined the items necessary to include and to exclude from the final list, in order to represent key items for evaluation of physical education teachers during observations. Once the lists of key items from the NASPE tool were discussed and agreed upon, the researchers sent this list to five experts in the field of physical education for validation. These five experts were chosen based upon the theory triangulation literature. Theory triangulation is the use of multiple perspectives to

interpret a single set of data from experts in various status levels within a field or profession (Guion, Diehl, & McDonald, 2011). The experts consisted of two professors of physical education, two doctoral candidates of physical education, and one K-12 teacher with over ten years of experience teaching physical education. The experts reviewed the list of key items and decided to either agree or disagree with the items as key indicators of teacher evaluation based on observations. For a disagreement from the team, a rationale for their decision was requested. Once the researchers received the list with feedback from the experts, they reconvened and discussed the outcomes. There was a consistent agreement amongst the experts (≥80%) on 45 items (81% from our initial decision), 36 to include and nine items not to include. After discussions, the researchers decided to exclude two items from the initial decision and to include one item as recommended by the experts. It was also determined that after feedback, seven key items would be combined into three items based upon overlapping definition (key items 12 & 54, 39 & 40, and 50, 51, & 52). It was determined that the finalized list of "key items" from the NASPE evaluation tool was 40.

Evaluation of four teacher evaluation instruments using NASPE key elements. Round two consisted of determining the presence of the 40 NASPE key items in each of the four teacher evaluation systems. The first step consisted of the researchers independently determining whether or not each of the 40 NASPE key items were present within the four teacher evaluation systems of inquiry. Once both researchers evaluated the four instruments for the 40 key items, multiple de-briefing sessions were held in discussing decisions, rationales, and overall findings for each individual evaluation instrument. After multiple discussions, the researchers agreed on the NASPE key items

presence on each teacher evaluation system. The second step in round two of data analysis was to validate research team member's findings. Thus, findings were sent out to three of the five aforementioned experts. The experts consisted of one professor of physical education, one doctoral candidate of physical education, and one K-12 teacher with over 10 years of experience teaching physical education. The experts reviewed the findings and determined if they agreed or disagreed with the researchers (see Table 1 for example). If agreed, the expert would check agreed, and if there was disagreement, the expert checked disagree and explained in a short rationale.

Table 1

Sample Researcher Findings for Expert Validation

NASPE			Teacher Evaluation System				
Domain	Key Items	Description	Domain	Item	Description	Agree/ Disagree	If Disagree please explain
Instruction	Lesson Introduction	Teacher provides an introduction that is appropriate for the lesson.	NA	NA			
Instruction	Learning Expectations/objec tives/instructional goals	Teacher states the skill or concept the students are to learn	Content	Content Accessibility	Discusses plan for making content accessible.		
Instruction	Content is accurate and current	Skills and content are taught accurately	Content	Conceptual Understand	Guides all Students to create…		

Note. NA signifies NASPE key item not found within teacher evaluation system of inquiry.

Once expert's data was received, the researchers discussed the agreement levels. The researchers held multiple de-briefing sessions to analyze the data and to finalize outcomes.

Trustworthiness of Data. Theory triangulation methods were used in analysis for rounds one and two of this study. Theory triangulation is the use of multiple perspectives from experts in various status levels within a field or profession to interpret a single set of

data (Guion, Diehl, & McDonald, 2011). Constant peer-debriefing was used for trustworthiness in determining credibility and dependability of the data. Peer-debriefing sessions are explained as "a process of exposing oneself to a disinterested peer in a manner paralleling an analytical session and for the purpose of exploring aspects of the inquiry that might otherwise remain only implicit within the inquirers mind" (Lincoln & Guba, 1985, p. 308). Member checks were also used to determine if experts agreed with the final items represented on each of the four teacher evaluation instruments that represented key items on the NASPE instrument. Lincoln and Guba (1985) explain member-checking as a technique for establishing the validity of an account (p. 298).

Results

Data from the following areas are displayed and discussed below: (a) the percentage of common parts and not applicable parts within each teacher evaluation system in comparison to the NASPE teacher evaluation tool; (b) key items from NASPE that are present in 100% of the four teacher evaluation systems; (c) key items from NASPE that are missing in ≥75% of the four teacher evaluation systems.

Danielson Framework for Teaching (FFT).

Common Parts and Not Applicable Parts. Results from the document analysis of the FFT indicate that seven items (17.5%) from the NASPE tool were not applicable within the FFT model, while 33 items (82.5%) were present (see Table 2).

Table 2

NASPE Physical Education Teacher Evaluation Tool VS. Danielson Framework for Teaching (FFT)

NASPE Domain	Key Items	Danielson Domain	Item

Instruction	Lesson introduction is appropriate	NA	NA
Instruction	Learning expectations/objectives/instructional goals are clearly communicated to students	Instruction Planning and Preparation	Communicating with students Setting instructional outcomes
Instruction	Content is accurate and current	Planning and Preparation	Demonstrating knowledge of content And pedagogy
Instruction	Content and tasks are developmentally appropriate and properly sequenced	Planning and Preparation Instruction	Designing coherent instruction Engaging students in learning
Instruction	Content and tasks are presented concisely and clearly, emphasizing key elements	Instruction	Communicating with students
Instruction	Engages students in learning by enabling all learners to participate through multiple modalities (opportunities to practice the skill).	Instruction	Engaging students in learning
Instruction	Opportunities for teachable moments are recognized and utilized	Instruction	Demonstrating flexibility and responsiveness
Instruction	Instruction is differentiated for all learners (accommodations and modifications are made for students with disabilities or varied learning styles).	Instruction	Demonstrating flexibility and responsiveness
Instruction	Specific, meaningful and timely feedback is provided to students (e.g., performance, efforts & positive contributions)	Instruction	Using assessment in instruction
Instruction	Content is linked to and promotes the transfer of learning within physical education units and among other subject content areas	Planning and Preparation	Setting instructional outcomes
Instruction	Student performance is continually assessed to guide instruction	Instruction	Using assessment in instruction
Instruction	Independent learning is promoted ,encouraged, and reinforced through daily assessments	Instruction	Using assessment in instruction
Instruction	Lesson pace is appropriate	Instruction	Engaging students in learning

Instruction	Appropriate closure is provided	NA	NA
Evidence of Student Learning	Assessment is based on mastery of learning expectations which are aligned with local, state and national standards	NA	NA
Evidence of Student Learning	There is ongoing formal and informal assessment	Instruction	Using assessment in instruction
Evidence of Student Learning	Assessment criteria is communicated to students	Instruction	Using assessment in instruction
Evidence of Student Learning	Multiple assessment strategies and tools are used (formative and summative) to monitor student learning	Instruction	Using assessment in instruction
Evidence of Student Learning	Students are able to articulate relevance and transfer of learning	NA	NA
Evidence of Student Learning	Student progress is documented in a retrievable record-keeping system	Professional responsibilities	Maintaining accurate records
Evidence of Student Learning	Student progress and achievement is communicated regularly to relevant stakeholders	Professional responsibilities	Communicating with families
Management/Organization	Lesson plans and curriculum are aligned w/ current local, state, and national standards	NA	NA
Management/Organization	Instructional area is safe, orderly, and supports learning activities	The classroom environment	Organizing physical space
Management/Organization	Adequate and developmentally appropriate equipment is accessible and utilized	The classroom environment	Organizing physical space
Management/Organization	Instructional support materials are utilized to enhance the lesson.	Planning and Preparation The classroom environment	Demonstrating knowledge of resources Organizing physical space
Management/Organization	Students understand and adhere to class rules, routines and behavioral expectations	The classroom environment	Managing student behavior
Management/Organization	Class routines maximize instructional time	The classroom environment	Managing classroom procedures
Management/Organization	There is a behavior management plan that is fair, firm, and equitable	The classroom environment	Managing student behavior

Management/Organization	Appropriate behaviors are reinforced consistently	The classroom environment	Managing student behavior
Management/Organization	Students are actively monitored and closely supervised using effective management strategies	The classroom environment	Managing student behavior
Management/Organization	Students are appropriately grouped	The classroom environment	Managing classroom procedures
Management/Organization	Effective and smooth transitions are apparent	The classroom environment	Managing classroom procedures
Management/Organization	Allocated time is used effectively and efficiently allowing students to remain focused on the lesson and task expectations.	The classroom environment	Managing classroom procedures
Management/Organization	Students are engaged in relevant, meaningful physical activity a minimum of 50-60 % of the instructional time.	NA	NA
Management/Organization	Accurate records are maintained	Professional responsibilities	Maintaining accurate records
Learning Climate	Lifelong physical activity and skillful movement are promoted	NA	NA
Learning Climate	There is a safe, secure, learning environment that promotes, success, appropriate risk taking, positive self-expression and enjoyment	The classroom environment	Organizing physical space
Learning Climate	High expectations for learning and behavior are evident	The classroom environment The classroom environment	Establishing a culture for learning Managing student behavior
Learning Climate	Climate of courtesy and respect is established	The classroom environment	Creating an environment of respect And rapport
Learning Climate	Students support the learning of others	The classroom environment	Creating an environment of respect And rapport

Note. NA= not applicable

There was evidence supporting that the FFT and NASPE tool were very common

across the instructional domains. Both systems support maximum student engagement

and the teacher's ability to demonstrate pedagogical content knowledge in the classroom.

One area of disconnect that emerged from the data was related to the structure of the

lessons. Moreover, while the NASPE tool targets an appropriate lesson introduction and closure, these key items were not present within the FFT system.

Evidence of Student learning was supported in both the NASPE and the FFT systems. Thus, both systems supported the use of ongoing formal assessments along with the use of multiple assessment strategies to monitor student learning. An area within the Evidence of Student Learning domain that was supported in the NASPE tool but not present within the FFT was assessment of student mastery of learning expectations aligned with national, state and local standards.

There was also evidence that both systems support the management and organization of a healthy learning environment. Moreover, terms such as on-task behavior, grouping and classroom procedures are commonly used across both of the systems. Furthermore, one area that is targeted in the NASPE tool that is missing from the FFT is the amount of time students are to be engaged in meaningful physical activity.

Teacher Advancement Program (TAP)

Common Parts and Not Applicable Parts. Results from the document analysis of the Teacher Advancement Program (TAP) indicate that 5 items (12.5%) from the NASPE tool were not applicable on the TAP system, while 35 items (87.5%) were present (see Table 3).

Table 3

NASPE Physical Education Teacher Evaluation Tool VS. The Teacher Advancement Program (TAP)

NASPE		TAP	
Domain	Key Items	Domain	Item

Instruction	Lesson introduction is appropriate	Instruction	Lesson Structure and Pacing
Instruction	Learning expectations/objectives/instructional goals are clearly communicated to students	Instruction	Standards and Objectives
Instruction	Content is accurate and current	Instruction	Presenting Instructional Content Teacher Content Knowledge
Instruction	Content and tasks are developmentally appropriate and properly sequenced	Instruction	Presenting Instructional Content
Instruction	Content and tasks are presented concisely and clearly, emphasizing key elements	Instruction	Presenting Instructional Content
Instruction	Engages students in learning by enabling all learners to participate through multiple modalities (Opportunities to practice the skill).	Instruction	Activities and Materials Teacher Knowledge of Students
Instruction	Opportunities for teachable moments are recognized and utilized	Instruction	Motivating Students
Instruction	Instruction is differentiated for all learners (accommodations and modifications are made for students with disabilities or varied learning styles).	Instruction	Teacher Knowledge of Students
Instruction	Specific, meaningful and timely feedback is provided to students (e.g., performance, efforts & positive contributions)	Instruction	Academic Feedback
Instruction	Content is linked to and promotes the transfer of learning within physical education units and among other subject content areas	Designing and Planning Instruction	Student Work
Instruction	Student performance is continually assessed to guide instruction	Instruction	Academic Feedback
Instruction	Independent learning is promoted ,encouraged, and reinforced through daily assessments	Instruction	Academic Feedback
Instruction	Lesson pace is appropriate	Instruction	Lesson Structure and Pacing

Instruction	Appropriate closure is provided	Instruction	Lesson Structure and Pacing Assessment
Evidence of Student Learning	Assessment is based on mastery of learning expectations which are aligned with local, state and national standards	Designing and Planning Instruction	Assessment
Evidence of Student Learning	There is ongoing formal and informal assessment	Designing and Planning Instruction	Assessment
Evidence of Student Learning	Assessment criteria is communicated to students	Designing and Planning Instruction	Assessment
Evidence of Student Learning	Multiple assessment strategies and tools are used (formative and summative) to monitor student learning	Designing and Planning Instruction	Assessment
Evidence of Student Learning	Students are able to articulate relevance and transfer of learning	Designing and Planning Instruction	Student Work
Evidence of Student Learning	Student progress is documented in a retrievable record-keeping system	NA	NA
Evidence of Student Learning	Student progress and achievement is communicated regularly to relevant stakeholders	NA	NA
Management/Organization	Lesson plans and curriculum are aligned w/ current local, state, and national standards	Designing and Planning Instruction	Instructional Plans
Management/Organization	Instructional area is safe, orderly, and supports learning activities	The Learning Environment	Environment
Management/Organization	Adequate and developmentally appropriate equipment is accessible and utilized	The Learning Environment	Environment
Management/Organization	Instructional support materials are utilized to enhance the lesson.	The Learning Environment	Environment
Management/Organization	Students understand and adhere to class rules, routines and behavioral expectations	The Learning Environment	Managing Student Behavior
Management/Organization	Class routines maximize instructional time	The Learning Environment	Managing Student Behavior
Management/Organization	There is a behavior management plan that is fair, firm, and equitable	The Learning Environment	Managing Student Behavior Respectful Culture

Management/Organization	Appropriate behaviors are reinforced consistently	The Learning Environment	Managing Student Behavior
Management/Organization	Students are actively monitored and closely supervised using effective management strategies	The Learning Environment	Managing Student Behavior
Management/Organization	Students are appropriately grouped	Instruction	Grouping Students
Management/Organization	Effective and smooth transitions are apparent	Instruction	Lesson Structure and Pacing
Management/Organization	Allocated time is used effectively and efficiently allowing students to remain focused on the lesson and task expectations.	Instruction	Lesson Structure and Pacing Expectations
Management/Organization	Students are engaged in relevant, meaningful physical activity a minimum of 50-60 % of the instructional time.	NA	NA
Management/Organization	Accurate records are maintained	NA	NA
Learning Climate	Lifelong physical activity and skillful movement are promoted	NA	NA
Learning Climate	There is a safe, secure, learning environment that promotes, success, appropriate risk taking, positive self-expression and enjoyment	The Learning Environment	Environment
Learning Climate	High expectations for learning and behavior are evident	The Learning Environment	Expectations
Learning Climate	Climate of courtesy and respect is established	The Learning Environment	Respectful Culture
Learning Climate	Students support the learning of others	The Learning Environment	Respectful Culture

Note. NA= not applicable

There was evidence supporting that the NASPE tool and TAP system are very common across all domains. Moreover, the data indicate that the NASPE tool is 100% common with the TAP instrument in the Instructional domain.

Two areas where there is disconnect between the two systems are keeping accurate records of student progression and the communication of progression to relevant

stakeholders. While NASPE emphasizes the documentation of student progress and communication to relevant stakeholders, the TAP system does not. Lastly, the TAP system does not indicate amount of time students are to be engaged in meaningful physical activity during class.

Marzano's Teacher Evaluation Model

Common Parts and Not Applicable Parts. Results from the document analysis of Marzano's Teacher Evaluation Model indicate that 15 items (37.5%) from the NASPE tool were not applicable, while 25 items (62.5%) were present (see Table 4).

Table 4

NASPE Physical Education Teacher Evaluation Tool VS. Marzano's Teacher Evaluation Model

NASPE		Marzano	
Domain	Key Items	Domain	Item
Instruction	Lesson introduction is appropriate	NA	NA
Instruction	Learning expectations/objectives/instructional goals are clearly communicated to students	Classroom Strategies and Behaviors	Providing Clear Learning Goals and Scales (Rubrics)
Instruction	Content is accurate and current	NA	NA
Instruction	Content and tasks are developmentally appropriate and properly sequenced	Classroom Strategies and Behaviors	Previewing New Content Chunking Content into "Digestible Bites"
Instruction	Content and tasks are presented concisely and clearly, emphasizing key elements	Classroom Strategies and Behaviors	Chunking Content into "Digestible Bites" Practicing Skills, Strategies, and Processes
Instruction	Engages students in learning by enabling all learners to participate through multiple modalities (opportunities to practice the skill).	Classroom Strategies and Behaviors	Practicing Skills, Strategies, and Processes

Instruction	Opportunities for teachable moments are recognized and utilized	Classroom Strategies and Behaviors	Elaborating on New Information
Instruction	Instruction is differentiated for all learners (accommodations and modifications are made for students with disabilities or varied learning styles).	NA	NA
Instruction	Specific, meaningful and timely feedback is provided to students (e.g., performance, efforts & positive contributions)	Classroom Strategies and Behaviors	Celebrating Success Reflecting on Learning
Instruction	Content is linked to and promotes the transfer of learning within physical education units and among other subject content areas	Classroom Strategies and Behaviors	Previewing New Content
Instruction	Student performance is continually assessed to guide instruction	NA	NA
Instruction	Independent learning is promoted ,encouraged, and reinforced through daily assessments	NA	NA
Instruction	Lesson pace is appropriate	Classroom Strategies and Behaviors	Maintaining a Lively Pace
Instruction	Appropriate closure is provided	NA	NA
Evidence of Student Learning	Assessment is based on mastery of learning expectations which are aligned with local, state and national standards	Planning and Preparing	Attention to Established Content Standards
Evidence of Student Learning	There is ongoing formal and informal assessment	Classroom Strategies and Behaviors	Tracking Student Progress Celebrating Success
Evidence of Student Learning	Assessment criteria is communicated to students	Classroom Strategies and Behaviors	Providing Clear Learning Goals and Scales (Rubrics)
Evidence of Student Learning	Multiple assessment strategies and tools are used (formative and summative) to monitor student learning	NA	NA

Evidence of Student Learning	Students are able to articulate relevance and transfer of learning	Classroom Strategies and Behaviors	Previewing New Content
Evidence of Student Learning	Student progress is documented in a retrievable record-keeping system	NA	NA
Evidence of Student Learning	Student progress and achievement is communicated regularly to relevant stakeholders	Collegiality and Professionalism	Promoting Positive Interactions with Students and Parents
Management/Organization	Lesson plans and curriculum are aligned w/ current local, state, and national standards	Planning and Preparing	Attention to Established Content Standards
Management/Organization	Instructional area is safe, orderly, and supports learning activities	Classroom Strategies and Behaviors	Organizing the Physical Layout of the Classroom
Management/Organization	Adequate and developmentally appropriate equipment is accessible and utilized	NA	NA
Management/Organization	Instructional support materials are utilized to enhance the lesson.	Planning and Preparing	Use of Available Traditional Resources
Management/Organization	Students understand and adhere to class rules, routines and behavioral expectations	Classroom Strategies and Behaviors	Establishing Classroom Rules and Procedures
Management/Organization	Class routines maximize instructional time	Classroom Strategies and Behaviors	Establishing Classroom Rules and Procedures
Management/Organization	There is a behavior management plan that is fair, firm, and equitable	Classroom Strategies and Behaviors	Establishing Classroom Rules and Procedures Applying Consequences for Lack of Adherence to Rules and Procedures
Management/Organization	Appropriate behaviors are reinforced consistently	Classroom Strategies and Behaviors	Applying Consequences for Lack of Adherence to Rules and Procedures Acknowledging Adherence to Rules and Procedures

Management/Organization	Students are actively monitored and closely supervised using effective management strategies	NA	NA
Management/Organization	Students are appropriately grouped	Classroom Strategies and Behaviors	Organizing Students to Interact with New Knowledge Organizing Students to Practice and Deepen Knowledge
Management/Organization	Effective and smooth transitions are apparent	Classroom Strategies and Behaviors	Maintaining a Lively Pace
Management/Organization	Allocated time is used effectively and efficiently allowing students to remain focused on the lesson and task expectations.	NA	NA
Management/Organization	Students are engaged in relevant, meaningful physical activity a minimum of 50-60 % of the instructional time.	NA	NA
Management/Organization	Accurate records are maintained	NA	NA
Learning Climate	Lifelong physical activity and skillful movement are promoted	Classroom Strategies and Behaviors	Organizing the Physical Layout of the Classroom
Learning Climate	There is a safe, secure, learning environment that promotes, success, appropriate risk taking, positive self-expression and enjoyment	NA	NA
Learning Climate	High expectations for learning and behavior are evident	NA	NA
Learning Climate	Climate of courtesy and respect is established	Classroom Strategies and Behaviors	Understanding Students' Interests and Backgrounds Using Verbal and Nonverbal Behaviors that Indicate Affection for Students

Learning Climate	Students support the learning of others	Classroom Strategies and Behaviors	Understanding Students' Interests and Backgrounds Using Verbal and Nonverbal Behaviors that Indicate Affection for Students

Note. NA= not applicable

There was evidence supporting that both the NASPE tool and the Marzano system emphasize the use of clear expectations, objectives, and instructional goals within the instructional and classroom strategies/behaviors domain. Moreover both systems support the use of developmentally appropriate learning task that are accessible for all learners. Furthermore, progressions, chunking, and proper feedback during lessons are important common parts of both instruments. While there are many commonalities within important areas of this domain, data suggests that NASPE key items are not present on 42.9% of Marzano's Instructional domain. Some key items from the NASPE instrument that are missing from the Marzano system are: (a) appropriate intro and closure of the lesson; (b) the use of accurate and current content; and (c) accommodations and modifications for students with disabilities.

NASPE's Evidence of Student Learning domain is common with Marzano. Moreover, evidence of its presence was found in 70% of the Marzano system. The Marzano system emphasizes clear learning goals and scales along with assessments that are linked to content standards. Furthermore, one area that NASPE supports that is not mentioned within the Marzano system is the use of multiple assessment strategies, along with documentation of student progress within a retrievable records system.

The NASPE Management/Organization key items are found to be common in 64.29% of the Marzano system. Moreover, both systems support establishing rules and

classroom procedures along with an organized classroom environment. The Marzano system does not emphasize active monitoring and supervision, or pinpoint the amount of time students are expected to be physically active during instructional time.

The NASPE Learning Climate key items are 60% common with the Marzano system. Thus, both systems are keen on promoting an environment of courtesy and respect. Furthermore, two NASPE key items that are absent from the Marzano system are the emphasis of a safe and secure learning environment, and high expectations for learning and behavior.

Rewarding Excellence in Instruction and Leadership (REIL)

Common Parts and Not Applicable Parts. Results from the document analysis of the Rewarding Excellence in Instruction and Leadership (REIL) teacher evaluation system indicate that 13 items (32.5%) from the NASPE tool were not applicable within the REIL teacher evaluation system, while 27 items (67.5%) were present (see Table 5).

Table 5

NASPE Physical Education Teacher Evaluation Tool VS. Rewarding Excellence in Instruction and Leadership (REIL)

NASPE		REIL	
Domain	Key Items	Domain	Item
Instruction	Lesson introduction is appropriate	NA	NA
Instruction	Learning expectations/objectives/instructional goals are clearly communicated to students	Content	Content accessibility
Instruction	Content is accurate and current	Content	Conceptual understanding
Instruction	Content and tasks are developmentally appropriate and properly sequenced	Content	Task analysis
Instruction	Content and tasks are presented concisely and clearly, emphasizing key elements	Content	Conceptual understanding

48

Instruction	Engages students in learning by enabling all learners to participate through multiple modalities (opportunities to practice the skill).	Instructional strategies	Monitor and adjust
Instruction	Opportunities for teachable moments are recognized and utilized	Instructional strategies	Monitor and adjust
Instruction	Instruction is differentiated for all learners (accommodations and modifications are made for students with disabilities or varied learning styles).	NA	NA
Instruction	Specific, meaningful and timely feedback is provided to students (e.g., performance, efforts & positive contributions)	Instructional strategies	Feedback (during the lesson)
Instruction	Content is linked to and promotes the transfer of learning within physical education units and among other subject content areas	Content	Connections to content
Instruction	Student performance is continually assessed to guide instruction	Formative assessment	Real-time assessment
Instruction	Independent learning is promoted ,encouraged, and reinforced through daily assessments	Formative assessment	Student progress
Instruction	Lesson pace is appropriate	NA	NA
Instruction	Appropriate closure is provided	NA	NA
Evidence of Student Learning	Assessment is based on mastery of learning expectations which are aligned with local, state and national standards	NA	NA
Evidence of Student Learning	There is ongoing formal and informal assessment	Formative assessment	Real-time assessment
Evidence of Student Learning	Assessment criteria is communicated to students	NA	NA
Evidence of Student Learning	Multiple assessment strategies and tools are used (formative and summative) to monitor student learning	Instructional strategies	Feedback (during the lesson)
Evidence of Student Learning	Students are able to articulate relevance and transfer of learning	Content	Connections to content

Evidence of Student Learning	Student progress is documented in a retrievable record-keeping system	NA	NA
Evidence of Student Learning	Student progress and achievement is communicated regularly to relevant stakeholders	Professional responsibilities	Communication with families
Management/Organization	Lesson plans and curriculum are aligned w/ current local, state, and national standards	NA	NA
Management/Organization	Instructional area is safe, orderly, and supports learning activities	NA	NA
Management/Organization	Adequate and developmentally appropriate equipment is accessible and utilized	NA	NA
Management/Organization	Instructional support materials are utilized to enhance the lesson.	NA	NA
Management/Organization	Students understand and adhere to class rules, routines and behavioral expectations	Learning community	Routines & procedures
Management/Organization	Class routines maximize instructional time	Learning community	Routines & procedures
Management/Organization	There is a behavior management plan that is fair, firm, and equitable	Learning community	Routines & procedures
Management/Organization	Appropriate behaviors are reinforced consistently	Learning community	Monitoring and responding to student behavior
Management/Organization	Students are actively monitored and closely supervised using effective management strategies	Learning community	Monitoring and responding to student behavior
Management/Organization	Students are appropriately grouped	Instructional strategies	Monitor and adjust
Management/Organization	Effective and smooth transitions are apparent	NA	NA
Management/Organization	Allocated time is used effectively and efficiently allowing students to remain focused on the lesson and task expectations.	Instructional strategies	Teacher role
Management/Organization	Students are engaged in relevant, meaningful physical activity a minimum of 50-60 % of the instructional time.	NA	NA

Management/Organization	Accurate records are maintained	Formative assessment	Student progress
Learning Climate	Lifelong physical activity and skillful movement are promoted	NA	NA
Learning Climate	There is a safe, secure, learning environment that promotes, success, appropriate risk taking, positive self-expression and enjoyment	Learning community	Monitoring and responding to student behavior
Learning Climate	High expectations for learning and behavior are evident	Learning community	Responsibility for learning
Learning Climate	Climate of courtesy and respect is established	Learning community	Relationships
Learning Climate	Students support the learning of others	Learning community	Relationships

Note. NA= not applicable

There was evidence supporting that the NASPE Instruction key items are 71.42% common with the REIL system. Moreover, both systems are supportive of the importance of conceptual understanding, modeling and adjusting instruction, and active assessment. Furthermore, an area missing from the REIL system that is emphasized as a key item within the NASPE tool are an appropriate introduction and closure to the lesson.

Within the NASPE domain of Evidence of Student Learning, both systems support the use of ongoing assessment and multiple assessment strategies. Furthermore, the REIL system is not applicable for the NASPE key items of assessing students based on mastery of learning expectations linked to content standards, communicating assessment criteria to students, and the documentation of student progress within a retrievable records keeping system.

The NASPE Management/Organization key items are common in 57.14% of the REIL system. Moreover, both systems include extensive items related to classroom rules and procedures, as well as consistently monitoring and reinforcing student behavior.

Furthermore, key Management/Organization items from NASPE that are missing from the REIL system (42.85%) are: (a) curriculum and lesson plans aligned with standards; (b) emphasis of a safe and organized learning environment; (c) accessibility of adequate and developmentally appropriate equipment; (d) the utilization of instructional support materials to enhance the lesson; (e)emphasis of effective and smooth transitions; and (f) emphasis on the amount of time students are engaged in meaningful physical activity during instructional time.

NASPE Tool Key Items Present in 100% of the Four Systems

Evidence supports that 15 key items (37.5%) from the NASPE tool are present in 100% of the four teacher evaluation systems (See Appendix B for complete List). Moreover, of the 15 total key items present from NASPE within all four systems, seven (46.66%) are from the Instructional Domain, one (6.66%) is from Evidence of Student Learning, five (35.71%) are from Management/Organization, and two (13.33%) are from Learning Climate.

NASPE Tool Key Items Missing from the Four Systems

Evidence supports that five key items (12.5%) from the NASPE tool are missing from ≥75% of the four teacher evaluation systems (See Appendix B for a complete list). The five NASPE key items are: (a) lesson introduction is appropriate; (b) appropriate closure is provided; (c) student progress is documented in a retrievable record-keeping system; (d) students are engaged in relevant, meaningful physical activity a minimum of 50-60% of instructional time; and (e) lifelong physical activity and skillful movement are promoted. .

Discussion

Results are consistent with the literature that suggests generic teacher evaluation systems can be used to evaluate teacher effectiveness across multiple subject areas, including physical education (Danielson, 1996; Danielson, 2007; Marzano Research Laboratory, 2011). Moreover, two systems had a high percentage of the NASPE tool key items present (TAP 87.5%, FFT 82.5%). The Marzano Model had the least number of NASPE key items present within its system (62.5%). Based on the variance in the presence of the NASPE key items, the suitability of these generic teacher evaluation systems may not address physical education.

Evaluators need to be trained and fluent in the subjects they observe. Brandt et al. (2007) discussed the importance of evaluators being trained and fluent in the subjects that they observed to avoid subjective and bias scoring. It was also stated that typically, district evaluation systems provide little guidance as to what to observe across subject areas (Brandt et al., 2007). These statements are relevant to this document analysis study in that the language across all four of the teacher evaluation systems of inquiry is not physical education specific. With the need for quality physical educators in every classroom, it is very important that physical education teachers are properly evaluated and measured on teacher effectiveness within their subject area. Furthermore, the researchers and the experts that contributed to this analysis are trained professionals within physical education and were able to connect the language within each of the four systems domains to the context of physical education. This is not the case with most evaluators that lack that specific training (Brandt et al., 2007).

It was determined that a high percentage of the NASPE tool key items are present within the four teacher evaluation systems (See Appendix B). Moreover, key items that are missing from the four teacher evaluation systems are in areas that are pertinent in delivering quality physical education lessons or specific to this content. Furthermore, the five NASPE key items left out of a majority of the four teacher evaluation systems are key concepts that are taught by physical education teacher education programs worldwide.

Strength and Limitations

An identifiable strength of this study was the knowledge and ability of the research team to identify and connect the language from the four teacher evaluation systems to physical education.

Two identifiable limitations of this study were the small sample of systems, and the lack of validity of the NASPE tool. There are many known teacher evaluation systems used throughout the country. The four systems examined in this study were limited to one county in one state in the Western U.S. There are no known validation studies associated with the NASPE physical education teacher evaluation tool.

<div align="center">

Conclusion

</div>

This study is one of the first to examine the degree of alignment between a physical education teacher evaluation system and four commonly used systems by school districts. Moreover, physical education as a school subject is not immune to the recent call for evidence of student learning and teacher effectiveness across all subject areas. Thus, more measures of accountability are in the near future. Finally, the call for more

training of school administrators/evaluators is essential in allowing physical education teachers the opportunity to receive a fair and valid evaluation.

This study can serve as a springboard for more research within the area of teacher evaluation in physical education. Furthermore, future implications may be the development of an instrument that could complement currently used systems, or stand alone to measure teacher effectiveness specific to physical education.

In conclusion, it appears that the TAP and FFT tools are appropriate for use in evaluating physical education teachers, contingent on at least two criteria. They include: (a) proper training having been completed by school administrators on the use of the evaluation tool; and (b) the evaluator having adequate physical education-specific Content Knowledge and Pedagogical Content Knowledge.

CHAPTER 4

ADMINISTRATORS PERCEPTIONS OF PHYSICAL EDUCATION TEACHER

EVALUATION

People's different experiences allow them to build different perspectives or

mental models of how things occur in the world (Starbuck & Milliken, 1988). Moreover,

these different mental models may act as perceptual filters that help to determine both

what we notice, and how it is interpreted (Starbuck & Milliken, 1988). Sense-making

theorists assume that people see what they know, and their actions are based on

experience from their past. To understand individual's perceptions further and how

administrators interpreted teacher evaluation specific to physical education within the

school as an organization, the Sense Making Theory was used as the framework for this

study.

With the absence of standardized testing in physical education classrooms

(NASPE, 2012), the sole measure of teacher performance relies on the teacher evaluation

process in place at the district level, as well as the expertise of the school level

administrator charged with conducting the evaluation. Generally, the school principal is

the instructional leader and holds the role for successfully guiding the professional

growth of the teachers within the school, and determining whether or not teachers will

keep their job (Millman, 1981). Protheroe (2002), stated that a well-executed evaluation

of teaching calls for the understanding of standards for student learning, an in-depth

understanding of what good teaching looks like in all classrooms across each subject

taught, and a strong ability to communicate and provide appropriate feedback. Moreover,

56

principals must have an in-depth knowledge of each teacher's performance, skills and areas of strength and weakness.

Although physical education teachers and the subject itself remain marginalized within the school setting (Sheehy, 2011), teacher evaluation systems remain an important measure in determining effective teaching, professional growth and promoting in student learner outcomes. With school administrators as the likely school site evaluators in physical education teacher evaluation, identifying their perceptions of the overall evaluation process can be important in determining principals' value orientation towards the subject itself, and their qualifications for conducting formal high-stakes based teacher evaluations.

Kersten and Israel (2005) examined K-8 principals' perceptions of teacher evaluation, and revealed an array of mixed results towards the topic. Principals reported that even though they invested a large portion of their day implementing the district required evaluation system, they felt that it had little impact on improving individual teacher instruction, or student achievement. Principals' attitudes toward teacher evaluation can affect the evaluation process in determining effective teacher practices (Tziner, Murphy, & Cleveland, 2001). Furthermore, when there is low motivation or incentive for teacher evaluation, it may become a cursory procedure with no appreciable impact on actual teaching practice. Thus, this may lead to a school culture in which teacher evaluation is not taken seriously and opportunity for instructional improvement is missing (Donaldson, 2009).

There are many different teacher evaluation systems in use across the country. Four commonly used teacher evaluation systems used are: a) teacher advancement

program (TAP) (National Institute for Excellence in Teaching, 2013); b) rewarding excellence in instruction and leadership (REIL) (Maricopa County Education Service Agency, 2013); c) the framework for teaching (FFT) (Danielson, 1996, 2007), and d) the Marzano teacher evaluation model (Marzano Research Labortory, 2013; Marzano, 2003).

The TAP program was created to assist with recruitment and retention of teachers in restructuring the evaluation and rewards system within the school (Little, 2009). The system links accountability with compensation by focusing on the following 4 key elements: (a) multiple career paths; (b) ongoing applied professional growth; (c) instructionally focused accountability; and (d) performance-based compensation. School districts adopt this system, and after a year, monetary consequences are put into place by the administrators (National Institute for Excellence in Teaching, 2013).

The REIL teacher evaluation system is similar to the TAP system in that it ties teacher pay to evaluation outcomes and student success. REIL is a five year initiative that was funded by the Teacher Incentive Fund through the US Department of Education, and is specific to Maricopa County in the state of Arizona (MCESA, 2013) REIL targets five critical elements: (a) rigorous, fair and transparent educator evaluations; (b) targeted professional learning; (c) tools for measuring student success; (d) establishment of multiple career pathways; and (e) sustainable, differential, performance-based compensation.

Danielson's FFT was created as a way to help in the improvement of teacher instruction (Danielson, 1996). The system has four domains that target different areas of teaching: (a) planning and preparation; (b) classroom environment; (c) instruction; and (d) professional responsibility. Within each of these four domains are 22 specific

performance components, which also include 76 specific task elements that are measured across a rating scale: (a) unsatisfactory; (b) basic; (c) proficient; and (d) distinguished.

The Marzano evaluation model was put together based upon Marzano's previous works associated with teaching areas and behaviors (Marzano Research Labortory, 2013). Marzano's model includes four domains that target different areas of teaching. They include: (a) classroom strategies and behaviors; (b) preparing and planning; (c) reflecting on teaching; and (d) collegiality and professionalism.

With teacher evaluation policies raising fundamental questions about what exactly effective teaching is and how it can be measured, the higher skilled the evaluators are in determining teacher behaviors across the many dimensions in teacher evaluation systems, may lead to more accurate scores determining teacher effectiveness (Kimball & Milanowski, 2009).

A major concern with teacher evaluation systems nationwide is that evaluators are failing to identify and remove low performing teachers. A recent report by a newspaper group in Illinois stated that 83 percent of the state's districts had never given a tenured teacher an unsatisfactory rating (Sartain, Stoelinga, & Krone, 2010). Moreover, Halverson and colleagues pointed out that teachers are able to identify when their evaluators lack pedagogical content knowledge within their particular subject area. Thus, teachers felt that their evaluators were not qualified to evaluate instructional content decisions. This lack of ability to make valid evaluations often led to very little critical feedback in written evaluation results (Halverson, Kelley, & Kimball, 2004).

Value-added models (VAMs) of teacher evaluation have become a popular method for measuring the value a teacher adds to student learning. VAMs use a complex

statistical formula to determine if teachers are contributing to a student's growth from year to year. Many districts are using these models to make high stakes decisions such as promotion, tenure, pay, and termination. Rothstein et al. (2010) stated that while value-added models contribute to stronger analyses of school progress, program influence, and increased validity of evaluations, these models alone are not reliable and valid indicators of teacher effectiveness. Most recently, value-added models have been criticized for assessing teachers based upon student outcomes, and not taking into consideration the countless variables outside of the classroom that contribute to student success (Amrein-Beardsley, & Collins, 2012; Berliner, 2014). The use of VAMs has the potential to affect physical education teachers, on the basis that physical education teachers are now starting to be held accountable for school wide standardized testing data.

There are no known studies to date that examine administrators' perceptions in evaluating physical educators. Therefore, the purpose of this study was to determine K-8 school administrators' perceptions of conducting formal teacher evaluation of physical education teachers in today's context of high stakes accountability approaches to teacher evaluation. The research question guiding this study was how do administrators' perceive the evaluation process specific to physical education? The emphasis of this question was based around four specific foci: (a) perceptions of the value of physical education; (b) perceptions and understanding of measures used to determine teacher effectiveness in physical education; (c) perceptions and understanding of measures used to determine student achievement in physical education; and (d) perceived ability to conduct teacher evaluation in physical education.

This study took a phenomenological approach to understanding administrators' perceptions to teacher evaluation specific to physical education. Locke, Silverman, and Spirduso (2010) noted that a phenomenological approach aims to understand the meaning of something from the vantage point of someone who actually experiences the phenomenon. Participants were asked to share descriptions, views and beliefs based upon their current and past experiences.

Methods

Participants and Setting

The participants in this study were administrators from one urban school district in the Western United States. The administrators were responsible for both the formative and summative evaluations of the Physical education teachers within their schools. Demographic information for the administrator is presented in Table 6.

Table 6

Administrator Interviewee Demographics

Administrator	Gender	Ethnicity	Experience	Taught Physical Education
Jack	M	Caucasian	8 Years	No
Barbara	F	Caucasian	Over 10 Years	No
Audrey	F	Hispanic	9 Years	Yes
Karen	F	Caucasian	Over 10 Years	No
Fred	M	Caucasian	8 Years	No
Eric	M	Caucasian	Over 10 Years	No
Paris	F	Caucasian	Over 10 Years	No
Pamela	F	Caucasian	Over 10 Years	No
Jill	F	Caucasian	3 Years	No
Susan	F	Hispanic	2 Years	No

Note. M=Male; F=Female.

The district is comprised of 19 elementary schools and 6 middle schools, and serve 17,756 students. Of those, 65.31% Caucasian, 15.33% Hispanic, 8.16% African American, 8.46% Asian/Pacific Islander, and 2.7% American Indian. ESL students make

up 2.9% of the district population. Of all students in the district, 23.4% were eligible for free and reduced lunch.

The selected district is "moving toward" using high-stakes teacher evaluation, but had not yet fully implemented it at the time of data collection.

As required by the University Institutional Review Board, each participant signed informed consent prior to the beginning of the study (see Appendix C).

Recruitment. Various districts were targeted for this study, but it is assumed that because of new developments in teacher evaluation protocols, this study was rejected. One district did agree to accept this study.

Pilot Study Protocol

A pilot study was conducted prior to the formal study that provided an opportunity for the researchers to fine-tune the interview protocol. Two administrators were chosen from outside the district and served as pilot study participants. The data collected during the pilot study was not used for the formal study.

Instrumentation

Informal survey. A survey was developed by the researcher to target the entire population of administrators in one district. The survey served three specific purposes: (1) gather demographic information; (2) reveal administrators level of agreement on statements concerning physical education teacher evaluation; and (3) sample administrators for formal semi-structured interview (see Appendix D for complete survey). Each statement in the survey was evaluated for internal consistency using Cronbach's alpha ($\alpha = .81$). Two content experts evaluated the items of this survey for content appropriateness.

Informal semi-structured interview. Interview questions were created by the researcher based on four specific foci: (a) perceptions of the value of physical education; (b) perceptions and understanding of measures used to determine teacher effectiveness in physical education; (c) perceptions and understanding of measures used to determine student achievement in physical education; and (d) perceived ability to conduct teacher evaluation in physical education. Two content experts evaluated the questions for content appropriateness (see Appendix E for complete interview guide).

Data Collection

Data from this study came from two sources. The first source was an informal survey created by the researcher. The survey was used to target the entire administrator population from the participating district to gain perceptions of physical education teacher evaluation, and as a source to sample from for the formal interviews. An email was sent out to all administrators (N=38) in one district containing a link to a short survey of 21 questions. Questions on the survey required answers using a Likert scale selection process (1=strongly disagree; 2=disagree; 3=agree; 4=strongly agree) (See Appendix D for complete survey). Of the 38 administrators that received the invitation to participate in the survey, 20 completed the survey (one administrator was removed for only having filled out demographic portion of survey). SurveyMonkey ® was used in administering the survey. All administrators who completed the survey were then invited to participate in a formal interview. Of the 19 administrators, 10 agreed to participate in the formal interview.

The second source of data was a formal semi-structured interview with 10 school administrators. The interview questions were created to reveal perceptions of teacher

evaluation specific to physical education. Interviews lasted approximately one hour each. Interviews were recorded with a digital voice recorder. Digital voice records were transcribed using Mac OS X Mountain Lion Dictation.

The questions used in the interviews were created by the research team members to fit the specific context of the study. Questions used in the interviews were grouped around four specific foci: (a) perceptions of the value of physical education; (b) perceptions and understanding of measures used to determine teacher effectiveness in physical education; (c) perceptions and understanding of measures used to determine student achievement in physical education; (d) perceived ability to conduct teacher evaluation in physical education. Not all questions listed were necessarily used in the formal study. Probing questions were used in leading up to the next chosen question (see Appendix E for a copy of the interview guide).

Data Analysis

Surveys. With the use of Microsoft Excel 2010, descriptive statistics were calculated for each question. In addition, for each survey statement, the percentage of respondents who scored it as "Agree" or "Strongly Agree" was calculated.

Interviews. Interview data were analyzed using constant comparison methods (Bogdan & Biklen, 2007). Common themes were identified and coded using the Dedoose (www.dedoose.com) online qualitative analysis software program (Dedoose Version 4.5, 2013). Two research team members independently reviewed all transcripts. Team members used frequent peer de-briefing sessions to determine that all themes had emerged from the data, to negotiate themes, and to minimize researcher bias (Lincoln & Guba, 1985, p. 308). Member checks were used in determining whether themes and

interpretations of participant statements were accurate (Kvale & Brinkman, 2009). Final transcripts and themes were sent to all participants' to ensure accuracy of findings. No changes were recommended by the administrators. Finally, a negative case analysis was utilized to ensure the accuracy of findings. Negative case analysis involved searching for and discussing any elements of the data that did not support or that appeared to contradict common themes that emerged from the data (Bogdan & Biklen, 2007).

Results

This study examined administrators' perceptions of teacher evaluation specific to physical education. The research question was; how do the administrators perceive the evaluation process specific to physical education? The emphasis was based around four specific foci; (a) perceptions of the value of physical education; (b) perceptions and understanding of measures used to determine teacher effectiveness in physical education; (c) perceptions and understanding of measures used to determine student achievement in physical education; and (d) perceived ability to conduct teacher evaluation in physical education. Findings of this study are presented by showing statistical data from the informal survey and discussing its meaning, along with discussing the common themes that emerged from the interview data.

Informal Survey Data

Descriptive statistics (*M, SD*) for survey statement response rates to Agree and Strongly Agree are presented in Table 7. Administrators' value physical education and find that it is just as important as core subject areas (e.g., Mathematics). Moreover, administrators believe that physical education teachers need to be held to the same high expectations as teachers of other subject areas. Furthermore, administrators find that

teacher evaluations in physical education are a useful tool for professional growth, as well as a reliable measure of teacher effectiveness.

Table 7

Administrators Perceptions of Physical Education Teacher Evaluation

Survey Statements	Mean	SD
1- Physical education is just as important to whole child development as are "core subjects" (e.g., Mathematics):	3.53	0.51
2- Teacher evaluation is as important for physical education teachers as it is for teachers of "core subjects":	3.84	0.37
3- Physical education teachers should be held to the same expectations as teachers of other school subjects:	3.68	0.48
4- Teacher evaluations are a useful tool for professional growth in physical education:	3.63	0.50
5-Teacher evaluations are a reliable measure of teacher effectiveness in physical education:	3.53	0.51
6-Current teacher evaluation practices impact teachers' classroom practices positively:	3.58	0.51
7-Evidence of student growth and achievement is (or "should be"??) an important factor of teacher evaluation in physical education:	3.37	0.68
8-My district's teacher evaluation system (e.g., Marzano) used within my district can accurately assess/determine the pedagogical content knowledge of physical education teachers:	3.00	0.67
9-I am skilled in accurately employing the current formal teacher evaluation tool (e.g., Marzano) when evaluating the performance of classroom teachers:	3.28	0.46
10-I am skilled in accurately employing the current formal teacher evaluation tool (e.g., Marzano) when evaluating the performance of physical education teachers:	3.00	0.59
11-I am highly skilled in using the data collected through the district's current teacher evaluation system for use in post-observation feedback sessions with my physical educator(s):	3.11	0.81

12-Feedback sessions/Post evaluation conferences are valuable to the professional growth of my physical education teacher(s):	3.63	0.50

Note. Likert Scale 1=Strongly Disagree;2=Disagree;3=Agree;4=Strongly Agree a) *n*=19

Statements 8, 10 and 11 received less than an 80% response rate of either Agree or Strongly Agree (See Figure 1 for percentages). All three of these statements focused on how the administrators perceived the current evaluation system used within their respective district. Statement 8 focused on whether or not the administrator felt the current evaluation system could accurately assess/determine the pedagogical content knowledge of physical education teachers. Although the response rate was high (>75%) for administrators agreeing with this statement, there were still 21.05% of the administrators that disagreed. Statement 10 focused on whether or not the administrator felt skilled at employing the current teacher evaluation system within a physical education classroom setting. There was evidence that 21.05% of the teachers either disagreed or strongly disagreed with this, meaning they did not feel skilled at employing teacher evaluation within a physical education classroom. The largest percent of administrators either disagreeing or strongly disagreeing occurred with statement 11(31.58%). This statement focused on understanding how administrators perceived their ability to use data collected from the formal evaluation, and give feedback to the physical education teacher during the post evaluation conference. With 100% of the administrators either agreeing or strongly agreeing (see question 12 in Figure 1) that the feedback session is important to the professional growth of the physical education teacher, statement 11 responses reflected a slight disconnect with how administrators can effectively use the evaluation data during the feedback session.
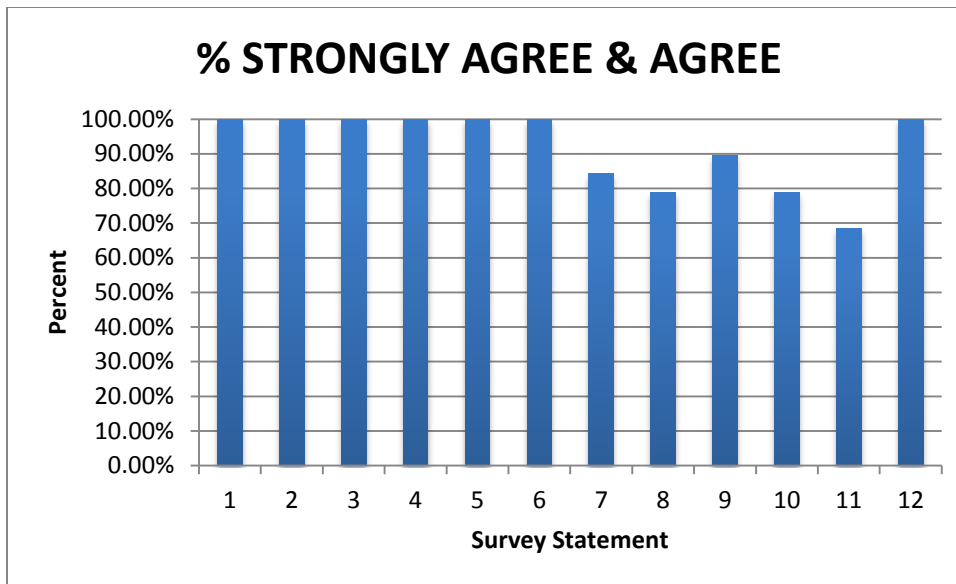
Figure 1. Percentages of Administrators that Agreed or Strongly Agreed with Survey Statement.

Interview Results

Four common themes emerged from the interview data; (1) "I value PE, but I live in reality" (administrators value physical education, but practice within their reality); (2) "good teaching is good teaching"; (3) "I know my limitations, and I want/need help" (relative to teacher evaluation in physical education); and (4) where's the training beef?

Theme 1: "I Value PE, but I Live in Reality". Administrators in this study value physical education. Moreover, the understanding that physical education was beneficial in the education of the whole child was very apparent. Furthermore, the administrators described the cognitive benefits physical activity contributes towards students' academic achievement. Karen has been an administrator for over 10 years. She stated that,

> I believe that physical education plays an important part, it plays a role in helping
>
> us accomplish everything else that we try to do with kids. Keeping the kids active
>
> and teaching them about healthy lifestyles and we see the benefit in the

68

classroom. I obviously can't control the amount of time spent in PE, but I can encourage brain breaks and other types of physical activity during the school day. I think that we see positive results in reading and math and everything else that they do.

Administrators in this study also identified the impact that physical education has on public health. They understand the obesity crisis that is affecting our youth due to sedentary lifestyles. Jill, who has been an administrator for three years, was very animated when asked if physical education was just as important as other subjects within her school; she stated,

> I do! Absolutely. Especially because our students, the kids, don't play outside as much as they should. They don't play outside as much as we used to. We know that there is a problem with childhood obesity, we see it. Here in school we see it. Kids are sitting more, playing video games more, on the computer all the time, watching TV. They spend the majority of the day inside. So absolutely, I feel that physical education is just as important as other subjects.

While administrators hold a strong value for physical education, they seem to understand that budget cuts and other factors controlled by policy makers are out of their control. Susan is relatively new to administration only having two years of experience. She stated that,

> To me I think it is very important. One, it is a state standard, it is part of the whole child, and the whole education process. I think it's necessary especially because the way we see the trends in everyone's physical fitness needs, so I think it's something that we need to have. Because of budget cuts kids are only able to

attend it once a week or once every six day cycle, depending on the school district.

Although administrators in this study valued physical education, they may lack knowledge as to what curricular models and instructional strategies are currently being used within their schools. Furthermore, when asked about curriculum in physical education, the administrators were unable to discuss the specifics of their schools physical education curriculum. Eric, who has over 10 years of experience, seemed a little uncertain as to what curriculum model was being used in his school. He stated,

> I know that elementary, they use the Pangrazi model. I am assuming that we use the same model in middle school, but please don't quote me on that one. I assume that the K-8 is planned together and that they use a lot of the same elements, but I know that they get into the more challenging competitive sports in the middle school program along with some of the skills, but to label it I am not sure if they use Pangrazi or not.

Audrey is a Hispanic female with nine years of experience. She has prior teaching experience in physical education. When asked about curriculum in her school she stated,

> Oh, um you know, that's an interesting question. I don't believe that there is, I mean, they, its not the Pangrazi, its not the dynamic physical education, it's kind of a morphed version of that, they do use part of that I would say a more morphed version of the dynamic. When you only have one person trained in the dynamic and we have a new teacher that's not trained in it. I would say its the old curricular model of fitness and games and skills. Yes fitness skills and games.

70

Quality physical education is another area where administrators seemed to have a lack of knowledge/understanding. Moreover, when asked about the definition of quality physical education, administrators seemed to relate their answer towards student participation and engagement levels. When asked about the definition of quality physical education, Barbara, who has over 10 years in the profession replied,

I don't think that quality physical education instruction is that much different than quality math instruction, it's about having an understanding of knowing what you want students to learn and do, and giving them lots of opportunities to practice doing that. So in PE that means giving students the opportunity to be moving and participating.

Jack who is an administrator with eight years of experience added when asked about quality physical education that,

Just like we want kids to be involved in their learning in the classrooms, its getting them active, its skill development, its practice and then application of that skill. Just like that circular model, were getting the kids to be aware of the skills involved in different activities, applying them and moving across the curriculum in different ways is the way I look at it.

Professional development (PD) is an area where the administrators had a lot of emphasis as far as promoting growth with their teachers. Moreover, administrators felt that PD was very important. However, they seemed to perceive PD for physical educators with an "us and them" approach, and at a loss to be of assistance to their school's physical educator. Thus, the administrator seemed to have a grasp on PD for core subject teachers, but had minimal input as to how much and what kinds of PD

physical education teachers received. Fred has been an administrator for eight years, and he responded about PD with the following statement,

> If it is, it's typically done at the district level. We don't provide it at the site level. Usually PE along with our other specialty areas if we're doing staff development days, we have one district staff development day in October. Typically, the PE teachers will get together and work on something and they will usually ask the principal's permission to go do this, and I always say absolutely, it's going to be more meaningful then working on our academic improvement plan. So I think anytime they can, they have to seek it out though and I know on Wednesdays which is our early release day, PE teachers will try get together and go over things. They will try to have departmental meetings among themselves and I think a lot of it is their own initiative in what they do, which is too bad, but that is just the way it works around here.

Jill who has three years of experience added when asked about PD,

> I don't know, but I think that our district will provide things for them. Two times a year we have our staff development days, and there are no students. Teachers are in professional development all day long, and often times PE teachers will go and do their own professional development and not be here at our site. So I know that they are offered professional development, but I know it's not a huge priority. I think a lot of times they have to seek it outside of the district.

Theme 2: Good Teaching is Good Teaching. As a group, the administrators believed that the evaluation system used within their district is intended to fit all content areas. Moreover, based upon common themes emerging from the data, they believed that

good teaching is good teaching and regardless of content area, teaching behaviors can be measured all the same. Jack has eight years of experience, and he stated that,

> PE just like in every other grade level or any other content area, I am not going
> to be the content expert, but even without that depth of knowledge, its again about
> the delivery of good instruction. Whether it be reading or writing or math,
> whatever it may be, we kind of look for the same things.

Paris who has over 10 years of experience added that,

> I think I am pretty comfortable and as an administrative group we are pretty
> comfortable that we can make this apply to any classroom whether it be pe, art,
> music because a lot of the elements, there is content, there is routine events
> happening constantly in the classroom.

> According to the administrators in this study, good teaching and teacher

effectiveness were measured by the engagement of the students. For instance, administrators appeared to gage the level of student activity as a primary indicator of both good teaching and teacher effectiveness. Thus, when asked about the measurement of teacher effectiveness during a formal observation, Barbara, with over 10 years of experience, stated that,

> "It's measured by the level of engagement the kids are engaged in."

Audrey, an administrator with nine years of experience also stated,

> "I would say that this is a hard one. I would say data based on the understanding
> of….I would say the level of engagement/the level of activity versus sitting time."

Karen, who had over 10 years teaching experience added,

I want to see that everyone is participating. So if there is an act or activity the

teacher is expecting everyone to do I will actually count how many kids are doing

it and divide it by the total number of kids. What I shoot for, and I think that this

is a reasonable goal is somewhere between 80% and 100% percent participating

100% of the time.

In addition to engagement levels of students as a measure of good teaching and

teacher effectiveness, administrators also looked at procedures such as safety and

classroom management as key areas. Susan, who has two years of experience stated,

Besides student engagement, another thing I would say would be is the

classroom organized, is it safe, are there sticks flying across the room, is a lesson

going smooth, are all the kids being monitored, are they getting feedback, are the

kids moving, is there talking going on, are the kids off task.

Paris, who has over 10 years of experience, when asked about measures of good

teaching and teacher effectiveness added,

"Classroom management in PE. It would look like this to me; is it safe, are

students being held accountable, it is optional whether or not the students are

participating."

While administrators believed that the evaluation system used within their district

was intended to fit all content areas, they also indicated that the current evaluation system

did lack subject matter specificity, making a fair assessment from someone without a

physical education background very difficult. Barbara (over 10 years) stated,

I don't have the knowledge or skill to really and truly evaluate my physical

education teacher in the way that someone with a PE background could do. I

74

don't try to cover that up either. I am pretty honest with my special area folks. If I

don't know I ask them to tell me what I am looking for.

Audrey (9 years) added,

I feel like this is our model, and now how are we going to make teachers try to fit

into this model? I have trouble with that anyway but for example, let's just say I

am going to choir, how are they going to provide an answer to question six

about test and hypothesis? How about cognitive complex and provide what those

two evidences are? You know I sometimes feel like we're trying to fit a square

peg into a round hole.

Theme 3: "I Know My Limitations, and I Want/Need Help".

Administrators in this study acknowledged their limitations relative to teacher

evaluation specific to physical education. They pointed out that their pedagogical content

knowledge may be lacking in the field of physical education and that more training

specific to the field as well as outside sources of expertise may be needed. Karen with

over 10 years of experience stated,

If I were a PE teacher I would want more specific feedback in my area but

realistically I don't know how possible that is without bringing in an expert in art,

or PE, or music to evaluate them. I think it would be very beneficial if at a future

principals meeting the district would have a physical education teacher come out

and talk to us about and explain to us exactly what we should be seeing at the

different grade levels and this is also the design of the format of our lessons so

that it would strengthen my knowledge of because I kind of feel like it's

rudimentary right now, I know that there is a scope and sequence.

Barbara (over 10 years) stated that,

> I would never want to give up the evaluation of the physical education teacher,
> but I wonder if there wasn't someone that came and looked at it really from that
> physical education standpoint about the quality of what's happening for kids as
> far as instruction, if there wasn't someone that could help do that.

As a group, the administrators stated the importance of the feedback sessions
based upon the importance of reflection. Moreover, administrators counted on feedback
sessions to assist in the professional development of the physical education teacher.
These sessions were often teacher led with the administrator relying on the expertise of
the teacher in the discussion of lesson results. Eric (over 10 years) stated,

> Why did you pull this group aside and have them do that, you know what I
> mean? So there is a lot for me, you know, my post evaluation conferences with
> my specialty areas are those types of questions because I want to get their level of
> thinking, because they have a level of thinking and expertise about their subject
> area that I just don't have because I haven't had those experiences. My classroom
> teachers, a lot of times I know why they did this or why they did that. I don't
> have to ask. I know why you pulled that kiddo, and I know why you did this. I
> don't particularly know that for PE and music. I'm like oh really, their feedback
> conferences is really about giving them the opportunity to talk about their
> rationale and chain of thinking and those types of things.

Jack (8 years) stated,

> It is really hard for us to dive deep into each subject area. Sometimes we hear
> that as a criticism, like hey, you don't know my content as well as I do. I

shouldn't. I shouldn't know your content as well as you do and if I do um, you are the grade level content expert, the content expert in your subject area.

Theme 4: Where's the Training Beef? As a group, administrators identified a weakness in the area of teacher evaluation training. Thus, administrators felt that there was a shortcoming within their administrative certification process in regards to how to conduct effective teacher evaluation. Furthermore, the administrators within this study felt they resorted to "on the job training" upon entering into their first leadership position. Paris with over 10 years of experience stated,

> I would say the administrative license share program didn't do jack for me in regards to [laughter] preparing me to be an instructional leader. I would hope that they're getting better. You get out in the field, and it's, kind of, like a sink or swim. That's where you'll get your experience.

Karen (over 10 years) stated,

> Well it paled in comparison to on-the-job training so maybe I had one three credit class that talked about supervision. The university did not prepare me for everything I needed to know about supervision and evaluation.

Audrey (9 years) added that,

> We received very little training during my master's program. And then, when ordered to become a qualified evaluator, we went to two day training. Now this model the Marzano, we've received extensive training. Well over 70 hours, very extensive training. So with this new model we received a lot of training, but when I went to become an administrator we didn't, we got little to none or very inadequate training.

77

Susan, who has been an administrator for 2 years added,

> My schooling for administration, I didn't receive any on teacher evaluation. I
> think that we took one class where we did research on them, but I don't remember
> ever covering them. It is interesting that I did not have any training during my
> school. You would think that you would get training on teacher evaluation. We
> did talk about the needs of looking at teachers and evaluating them, but not
> actually like looking at different tools and comparing them or learning how to use
> them.

Discussion

This study was conducted to develop deeper understanding of administrators'
perceptions of physical education teacher evaluation. The four common themes that
emerged from the data will be discussed within this section.

"I Value PE, But Live in Reality"

The good news is that evidence from this study supports that the administrators as
a group valued physical education and understand the benefits it has on students. This
aligns with the findings from Sallis, McKenzie, Kolody, and Curtis (1996), who noted
that a high percentage of school administrators believe that physical education can
enhance concentration, decrease discipline problems, and improve academic
performance. Hence, one would think that with a high percentage of administrators that
value the subject, it would be a higher priority within the curriculum. Furthermore, The
No Child Left Behind Act (2002) has played a huge role in the reduction of physical
education in the schools, making it a low priority within the curriculum (Barosso,
McCullum-Gomez, Hoelscher, Kelder, & Murray, 2005; Prince et al., 2008).

78

Administrators understand that even though they value physical education, their hands are tied due to national, state, and district policies, and the constant pressure to improve students' academic achievement scores, along with increasingly prevalent use of high-stakes teacher evaluation practices (Amrein-Beardsley & Collins, 2012; Berliner, 2014; Herilhy, 2012; Hill, Kapitula, & Umland, 2011).

Administrators defined quality physical education in their own words, and all but one administrator was unfamiliar about which curricular model was used in their schools. This evidence suggest that there is a major disconnect with the reality of what is seen by these administrators, and what is supposed to be seen in a quality physical education setting. These results are consistent with studies that found a high percentage of principals are unaware and far removed from day to day realities in physical education (Locke, 1975; Lounsbery, McKenzie, Trost, & Smith, 2011). This poses a major problem as the delivery of quality physical education is vital for enhancing physical activity opportunities, benefitting overall student health, and developing skills, attributes and behaviors to be active for life (Le Masurier, & Corbin, 2006; McKenzie, Marshall, Sallis, & Conway, 2000). Furthermore, quality physical education programs can have positive outcomes on student achievement and overall well-being (Coe, Pivarnik, Womack, Reeves, & Malina, 2006). Thus, the Centers for Disease Control and Prevention (2009) have recognized school physical education as an important and available resource for promoting physical activity and healthy behaviors amongst children. Finally, it is very important that the principal, as the school curriculum leader, is aware of current curriculum trends within physical education. This is vital in knowing that students are receiving a quality physical education (San Diego State, 2007).

79

Professional Development (PD) was identified as being very important in the growth of the physical education teachers. However, administrators approached PD with an "us vs them". Moreover, administrators had a strong grasp on the types and amount of PD for core subject teachers, but when asked about PD for physical educators, they were unable to provide the same information about their physical education teachers. This is quite concerning, as professional development has been linked to both teacher development and student learning (Huffman & Thomas, 2003). Wang and Ha (2008) pointed out that a major issue related to professional development for physical education teachers is the lack of support from the school site principal. Moreover, it is suggested that multiple stakeholders are to be involved with the professional development of the physical education teacher if it were to be effective, and that it should be considered through multiple lenses and aligned with elements such as district policies and curriculum requirements (Wang & Ha, 2008).

"Good Teaching is Good Teaching"

Administrators were convinced that the teacher evaluation system used within their district was applicable to all content areas, including physical education. The term "good teaching is good teaching" was used quite often during the interview process. Thus, it is assumed that the same teaching behaviors that are observable within a classroom are observable within the realm of a physical education setting (e.g. gymnasium, playing field). Locke (1975) explains the complexities of the physical education classroom and discussed the major differences it has from those present in the classroom. Moreover, students in physical education are active and mobile, sometimes within areas spanning an acre. Furthermore, students are not confined to a desk, making

80

classroom management and direct instruction more challenging. Finally, while teaching behaviors may be defined the same across all subject matters, the ecology of a physical education setting and how those behaviors may look are different (Locke, 1975).

Even though the administrators in this study felt that the evaluation system used within their schools was applicable to all content areas based upon a "good teaching is good teaching" philosophy, they all did agree that the system lacked content matter specificity, making a fair assessment of the physical education teacher more difficult. This outcome is similar to what Kimball (2002) found in which administrators felt comfortable giving generic forms of feedback on various teaching strategies, but lacked content knowledge, making a content related evaluation difficult.

"I know My Limitations and I Want/Need Help"

Administrators acknowledged that there is a lot of room for improvement in regards to the evaluation of physical education teachers. The administrators agreed that they do in fact lack content knowledge, thus making valid assessments of teaching performance in physical education very difficult. These outcomes are consistent with the literature that points out the lack of content knowledge that administrators face when evaluating teachers in not only physical education, but in all subject areas (Donaldson, 2009; Halverson, Kelley, & Kimball, 2004; Kimball, 2002). Related to the lacking content knowledge administrators felt they possess, they all are in agreement that more training specific to physical education content and instructional methods would be beneficial in conducting more effective teacher evaluations.

Feedback is considered a very important component of the teacher evaluation process (Danielson, 1996). Administrators in this study agreed that the feedback sessions

81

are very beneficial regardless of the level of content knowledge they have.

Administrators pointed out that these sessions are often teacher led. Moreover, this goes back to the fact that there is a lack of pedagogical content knowledge from the administrator's standpoint. With feedback and reflection being a major component in the professional growth of teachers, the ability for the administrator to discuss instructional strategies based on specific content is vital.

Prior to becoming an administrator, there is a certification process mandatory to fulfilling the position. Data from this study suggest that administrators feel they did not receive adequate training to conduct teacher evaluation during their certification program. These outcomes are consistent with studies that have evidence of principal candidates and existing principals being ill-prepared and inadequately supported to act as curriculum leaders, while fulfilling all other demands of the job (Levine, 2005; Peterson, 2001). Moreover, it was agreed upon that on the job training was where they received the bulk of their professional development in regards to most procedural responsibilities, to include teacher evaluation. Donaldson (2009) discussed the internal constraints to differentiation in teacher evaluation pointing out that "without high-quality professional development, evaluators will not evaluate accurately and the evaluation will likely have little impact on teaching or learning" (p.9).

Strengths and Limitations of the Study

An identifiable strength of this study is that it is one of the first to examine administrators' perceptions of teacher evaluation specific to physical education. This study may be groundbreaking and encourage future research in the area of teacher evaluation specific to physical education. Moreover, with the direction in which

82

accountability measures are headed for non-tested subjects, it is important that more research is conducted in this area.

There are three identifiable limitations within this study: (a) small sample size; (b) limited number of interviews; and (c) specificity to one school district. This study had a small sample size, and all participants were from one school district. This may have caused a lack of generalizability to the larger population (Locke, Silverman, & Spriduso, 2010). Furthermore, having one interview inhibited the opportunity for follow up questions which may have reduced the chances for richer data (Patel, & Doku, & Tennakoon, 2011).

Conclusion

School administrators have the responsibility of determining the effectiveness of physical education teachers within our public schools. Administrators' understanding the importance of physical education is just not enough. They need to be advocates and supporters of quality physical education programs and understand what's going on within their schools physical education curriculum. We are living in an epidemic where children are suffering from obesity, diabetes, and other disease brought on by sedentary and unhealthy lifestyles. Having effective physical education teachers teaching our children is vital in combating this epidemic. Moreover, administrators are key players in determining physical education teacher effectiveness, as they are the ones evaluating and making high stakes decisions. Thus, the need for them to receive the proper training and education on effective physical education classroom practices, new trends in physical education curricula, and an understanding of the ecology in a physical education setting is vital in determining these high stakes decisions. Finally, administrators need to be given

the proper instrumentation to give a valid evaluation to physical education teachers. While generic systems may suffice in the eyes of policymakers and stakeholders for now, future instrumentation must reflect what physical education teachers and evaluators consider being the essential elements of instruction in a physical education setting and yielding reliable results.

This study is one of the first to examine administrator's perceptions of teacher evaluation specific to physical education. Evidence from this study suggests more training for administrators in physical education teacher evaluation, along with a content specific evaluation instrument are greatly needed. More research will be necessary in the area of physical education teacher evaluation, as the call for evidence based outcomes in teacher effectiveness and student achievement are brought to the table.

CHAPTER 5

PHYSICAL EDUCATION TEACHER PERCEPTIONS OF TEACHER EVALUATION

Swanson (1999) defined performance as the way something or someone

functions. Performance can negatively or positively affect a single person, other people,

or an entire organization (Swanson, 1999). Moreover, schools are organizations that are

committed to performance, including human resource development, management, and

quality improvement linked to teacher evaluation. Therefore, to further understand

physical education teacher evaluation systems, the Performance Improvement Theory

was used as a framework for this study (PIT) (Swanson, 1999). The PIT frameworks

main focus is for people to learn and grow. When this occurs, people are empowered to

create results and make a difference (Swanson, 1999).

The No Child Left Behind Act of 2002, and most recently President Obamas Race

to the Top in 2009, have caused stakeholders in education to raise their interest in teacher

accountability. Because of the elevated accountability and higher emphasis on showing

evidence of teacher effectiveness and student achievement, various performance based

and value-added models of teacher evaluation have been adopted by school districts

(Little, Goe, & Bell, 2009).

Performance based teacher evaluation systems are meant to measure teacher

effectiveness in the classroom by using multiple rating categories across general bodies

of knowledge and skills for teaching (Shakman et al, 2012). Two commonly used

performance based teacher evaluation systems are the Marzano teacher evaluation system

(Marzano, 2003), and the Danielson framework for teaching (Danielson, 1996). These

systems both use a series of domains that target different areas of teaching (e.g. planning

and preparation, classroom environment). Moreover, within these domains are specific task elements that are measured across different rating scales. These systems are typically used by school administrators during formal classroom observations. Furthermore, outcomes from these evaluations are used to rate teachers effectiveness based on proficiency levels scored by the evaluator (Little, Goe, & Bell, 2009). There is evidence that suggests principals receive little training with these systems, and that scoring outcomes are based upon subjectivity (Brandt, Mathers, Oliva, Brown-Sims, & Hess, 2007; Jacob, & Lefgren, 2008).

Value-added models are now a widely adopted resource used by districts to measure the value a teacher adds to student achievement from year to year (Little, Goe, & Bell, 2009). One of the most widely used value-added models is the SAS Education Value Added Assessment System (Amrein-Beardsley, & Collins, 2012). The SASEVAAS claims to provide valuable diagnostics of students 'growth over time, and tie that growth back to the teacher. Moreover, these diagnostics are used to measure a teachers' effectiveness in the classroom (Little et al., 2009). Rothstein et al. (2010) noted that while value-added models (VAMs) contribute to stronger analyses of school progress, program influences, and increased validity of evaluations, these methods alone are not reliable and valid indicators of teacher effectiveness. Herlihy (2012) examined state and local efforts to investigate validity and reliability of scores from teacher evaluation systems. Few states seemed to be considering the negative unintended consequences of systems that may generally be perceived as being arbitrary. Furthermore, administrators from many states understand the importance of implementing better systems, but lack knowledge of implications for validity and

reliability of scores produced by their systems (Herlihy, 2012). Moreover, Berliner

(2014) pointed out the many exogenous variables associated with VAM's, indicating that

the major problem for value added approaches is assessing teachers based upon student

outcomes, when countless variables beyond the classroom affect achievement inside the

classroom.

Current evaluation systems often lack alignment with subject matter curricular

standards (Jerald, 2009). With so many different variations of evaluation system practices

(Little, 2009), most district level teacher evaluation policies provide little guidance on

what criteria to follow when observing, along with how to use and share feedback from

the evaluation process (Brandt, Mathers, Oliva, Brown-Sims, & Hess, 2007). These

inaccuracies, coupled with lack of support and insufficient training may affect school

cultures in which both administrators and teachers alike struggle with the process, do not

take the evaluation process seriously, or see little practical value (Donaldson, 2009).

There is evidence that traditional teacher evaluation programs have been based on

competing conceptions of teaching, often characterized by inaccuracies, insufficient

training and a lack of support (Darling-Hammond, Wise, & Kline, 1999; Loup, Garland,

Ellet, & Rugutt, 1996; Peterson, 1995). Moreover, Halverson, Kelly and Kimball (2007)

found that some teachers felt their evaluators lacked in pedagogical content knowledge

and were not qualified to evaluate on instructional content decisions. Zimmerman and

colleagues (2003) found that teachers viewed their evaluators as just going through the

motions and not showing much effort being put into teacher evaluation. Furthermore,

these actions gave mixed feelings as to whether or not the evaluation system was tailored

towards them or just a generic approach (Zimmerman, 2003). There is evidence that

teachers value communication and trust between themselves and the evaluator (Davis, 1988; Valentine, 1992). Moreover, teachers have shown interest in receiving "feedback" in one form or another.

Similar to music and art, physical education remains a marginalized subject receiving low priority and concern within school curriculum (Prince et al., 2008). With the obesity epidemic plaguing our nation and school physical education playing a key role in counteracting this dilemma (Pate, Davis, Robinson, Stone, McKenzie, & Young, 2006) , it is more important now than ever before that physical education teachers are receiving quality evaluations and given opportunity for professional growth and development.

There are no known studies that examine how physical educators perceive current teacher evaluation systems. Therefore, the purpose of this study was to develop a better understanding of current physical education teachers' perceptions of teacher evaluation systems. The research question for this study was; how does the physical education teacher perceive teacher evaluation? The emphasis of this question is based around 5 specific foci: (a) perceptions of the value of physical education within the school; (b) understanding of quality and standards based physical education; (c) perceptions of teacher evaluation in physical education; (d) confidence of the evaluator; (e) evidence of student learning.

This study takes a phenomenological approach to understanding physical education teachers' perceptions to teacher evaluation specific to physical education. Locke, Silverman, and Spirduso (2010) noted that a phenomenological approach aims to understand the meaning of something from the vantage point of someone who actually

experiences the phenomenon. Participants were asked to share descriptions, views and beliefs based upon their current and past experiences.

Methods

Participants and Setting

The participants in this study were physical education teachers from one urban school district in the Western United States. The teachers experience ranged from 1 to over 10 years. There were two sources of data collection used within this study; an informal survey, and formal semi-structured interviews. Demographics for both samples will be listed starting with the survey participants. Of the 22 participants that fully completed the survey, there were 11 males and 11 females. There were 21 teachers who indicated that their ethnic background was Caucasian and one that indicated as bi-racial background. For years of experience, two teachers had 1-5 years of experience, three had 5-10, and 17 had over 10 years of experience.

The demographics of the 10 teachers who volunteered to be interviewed are presented in Table 8. As required by the University Institutional Review Board, each participant signed consent prior to beginning the study (See Appendix C).

Table 8

Teacher Interview Demographics

Teacher	Gender	Ethnicity	Experience	Grade Level
Jennifer	F	Caucasian	9 Years	K-5
Tracy	F	Caucasian	Over 10 Years	6-8
Kevin	M	Caucasian	Over 10 Years	6-8
Elizabeth	F	Bi-Racial	1 Year	K-5
Kelly	F	Caucasian	Over 10 Years	K-5
Marie	F	Caucasian	Over 10 Years	K-5
Brittany	F	Caucasian	Over 10 Years	K-5
Johnny	M	Caucasian	Over 10 Years	6-8
Sharon	F	Caucasian	8 Years	K-5
Sheila	F	Caucasian	Over 10 Years	6-8

The district is comprised of 25 schools. There are 19 elementary schools and 6 middle schools. The district serves17,756 students. 65.31% Caucasian, 15.33% Hispanic, 8.16% African American, 8.46% Asian/Pacific Islander, and 2.7% American Indian. ESL students make up 2.9% of the district population. Of all students in the district, 23.4% are eligible for free and reduced lunch.

The selected district is "moving toward" using high-stakes teacher evaluation, but had not yet fully implemented it at the time of data collection.

Pilot Study Protocol

A pilot study was conducted prior to the formal study that provided an opportunity to the researcher to fine-tune the interview protocol. Two physical education teachers from outside the district were recruited to serve as pilot study participants. The data collected during the pilot study were not used for this study.

Instrumentation

Informal survey. A survey was developed by the researcher to target the entire population of physical education teachers in one district. The survey served three specific purposes: (1) gather demographic information; (2) reveal physical education teachers level of agreement on statements concerning physical education teacher evaluation; and (3) sample physical education teachers for formal semi-structured interview (see Appendix F for complete survey). Each statement in the survey was evaluated for internal consistency using Cronbach's alpha ($\alpha =.79$). Two content experts evaluated the items of this survey for content appropriateness.

Informal semi-structured interview. Interview questions were created by the researcher based on five specific foci: (a) perceptions of the value of physical education within the school; (b) understanding of quality and standards based physical education; (c) perceptions of current teacher evaluation practices in physical education; (d) confidence of the evaluator; (e) evidence of student learning. Two content experts evaluated the questions for content appropriateness (see Appendix G for complete interview guide).

Data Collection

Data from this study came from two sources. The first source was an informal survey created by the researcher. The survey was used to target the entire physical education teacher population from the participating district to gain perceptions of physical education teacher evaluation, and as a source to sample from for formal interviews. An email was sent out to all physical education teachers (*N*=33) in one district containing a link to a short survey of 18 questions. Questions on the survey required answers using a Likert scale selection process (1=strongly disagree; 2=disagree; 3=agree; 4=strongly agree) (See Appendix F). Of the 33 physical education teachers that received the invitation to participate in the survey, 22 completed the survey (2 teachers were removed for only having filled out demographic portion of survey). SurveyMonkey® was used in administering the survey. Teachers that completed the survey were invited to participate in a formal interview. Of the 22 physical education teachers that completed the survey, 10 agreed to participate in a formal interview.

The second source of data was a formal semi-structured interview with 10 physical education teachers. The interview questions were created to reveal their

91

perceptions of their districts current high-stakes teacher evaluation practices in general, and specific to the physical education context. Interviews lasted approximately one hour each. Interviews were recorded with a digital voice recorder. The researcher, using Mac OS X Mountain Lion Dictation, then transcribed digital voice records.

The questions used in the interviews were created by two research team members to fit the specific context of the study (see Appendix G for a copy of the interview guide).

Data Analysis

Surveys. With the use of Microsoft Excel 2010, descriptive statistics were calculated for each question. In addition, for each survey statement, the percentage of respondents who scored it as "Agree" or "Strongly Agree" was calculated.

Interviews. Interview data were analyzed using constant comparison methods (Bogdan & Biklen, 2007). Common themes were identified and coded using the Dedoose (www.dedoose.com) online qualitative analysis software program (Dedoose Version 4.5, 2013). Two research team members independently reviewed all transcripts. Team members used frequent peer de-briefing sessions to determine that all themes had emerged from the data, to negotiate themes, and to minimize researcher bias (Lincoln & Guba, 1985, p. 308). Member checks were used in determining whether themes and interpretations of participant statements were accurate (Kvale & Brinkman, 2009). This was done by sending all final transcripts and themes to participants' to ensure accurate statements. No changes were recommended by the teachers. Finally, a negative case analysis was utilized to ensure the accuracy of findings. Negative case analysis involved searching for and discussing any elements of the data that did not support or that

appeared to contradict common themes that emerged from the data (Bogdan & Biklen, 2007).

Results

This study examined physical education teachers' perceptions of teacher evaluation specific to physical education. Findings of this study are presented by showing statistical data from the informal survey and discussing its meaning, along with discussing the common themes that emerged from the interview data.

Informal Survey Results

Descriptive statistics (*M, SD*) for individual survey statements as well as the percentage of respondents who scored each statement with "Agree" or "Strongly Agree" are presented in Table 9.

Table 9

Physical Education Teachers Perceptions of Physical Education Teacher Evaluation

Survey Statements	Mean	SD
1-Physical education is a priority in my school curriculum:	2.59	0.67
2- Teacher evaluation is necessary in physical education:	3.41	0.50
3- Physical Education teachers need to be held to the same expectations as teachers of other school subjects:	2.95	0.84
4- Teacher evaluations are a useful tool for professional growth in physical education:	3.18	0.59
5- Teacher evaluations are a reliable measure of teacher effectiveness in physical education :	2.71	0.85
6- Teacher evaluation does have an effect on teacher practices:	3.05	0.67
7- Evidence of student growth and achievement is an important	2.71	0.85

factor of teacher evaluation in physical education:

8- I completely understand the current teacher evaluation system used within my district and know exactly what is expected of me:	2.73	0.83
9- I am confident that my administrator/evaluator is able to determine my effectiveness as a physical educator:	2.75	0.85
10- Feedback sessions/Post evaluation conferences are valuable to my growth as a physical education teacher:	3.14	0.64

Note. Likert Scale ranged from: 1=Strongly Disagree; 2=Disagree;3=Agree;4=Strongly Agree. a) *n*=22

All of the physical education teachers either agreed or strongly agreed that teacher evaluation in physical education was necessary. They also firmly believed that teacher evaluation was important for their professional growth, as evidenced in the responses to statements 4 and 10.

Some areas of the survey that received low support from the physical education teachers were statements 1, 7, 8, and 9. Only 50% of the respondents viewed that their program was a school priority (Statement 1). A high percentage of the physical education teachers did not believe that evidence of student growth and achievement were important factors of teacher evaluation in physical education. Statements 8 and 9 directly reflected the physical education teachers confidence in both the evaluation system and the evaluators ability to determine the effectiveness of the physical educator. Just fewer than 60% of the physical educators completely understood the current system being used in their district. Finally, there was evidence that physical education teachers were not very confident that their evaluators can determine their effectiveness as a physical educator.
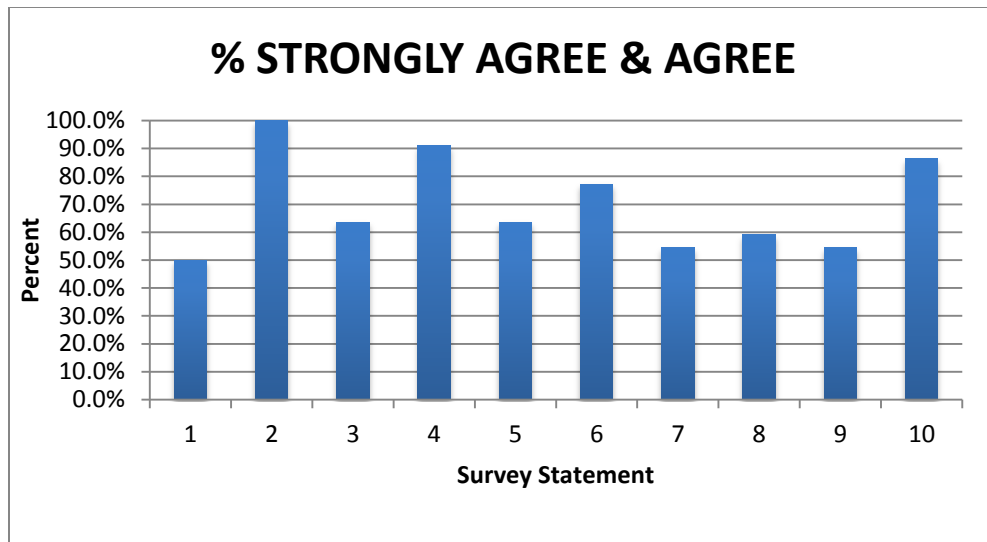
Figure 2. Shows percentages of physical education teachers that Strongly Agreed or Agreed with the statement.

Interviews

Results determined three themes for physical education teacher's views of the teacher evaluation process: (a) valued, but not prioritized?; (b) teacher evaluation in physical education is "greatly needed, yet not transparent; (c) "who do you trust" (physical educators are not confident in their evaluators).

Valued, but not prioritized. Physical educators feel that physical education is valued within their schools, yet not valued enough to make it a priority. Moreover, as a group the teachers claimed that physical education priority was based on the value that the school administrator placed on physical education. Tracy is a physical education teacher with over 10 years of experience at both the elementary and secondary levels. She explained that from her current and past experiences, it really depends on how the school principal feels personally about physical education. Moreover, if the principal shows value towards physical education, then the rest of the faculty will follow suit. Tracy stated that,

95

I think if they view it (principal) as not important, then the teachers view it that

way too. They'll pull kids out whenever they want, whereas we want them in

our classes working hard, learning, team building, whatever we're doing. If

everybody thinks they can pull the kids whenever they want, our program isn't

valued. Having an administration that values that makes a big difference.

When asked about the value and priority of physical education at her school, Marie who

has over 10 years of experience stated,

Well, I do feel like my principal feels it's very important, but, unfortunately,

when our district ran into money trouble and we had to eliminate one of the days,

I feel like priority-wise it did go down. We're equal to other special areas now,

and I'm not saying that that's bad or good, but, just time-wise, I think that we

have reduced a little bit. They've put so much emphasis on testing and the

standards and all that kind of thing now that I think, unfortunately, we are sort of

like a second step.

Brittany, who has over 10 years of experience, referred to the overall value of physical

education in her school to that of a grading system. She felt that value held a grade of a

B, but very low in priority. She stated,

Well, I'd say it's—if you rated them A, B, C, I'd say I'd probably give it a B as

far as the overall feeling, the general—the way people value it. I think they value

what we do. It's just I think that it's not as important as math and reading and

those things to, I would say, most people. I feel that they believe it is not as

important.

96

The physical education teachers felt that there was an "us and them" mentality within the culture of their school. Physical educators grouped themselves with other "specialty areas". Johnny, with over 10 years of experience explained his feelings with the following statement:

> I just think a lot of people don't have the same level of respect for physical education that I do. I think there's a lot of pressure on administration to—with the test scores and that's what a lot of parents are interested in, seeing those higher test scores. To get that, I think they try and cram as much academics as they can, and focus groups, and ways to improve kids' level in reading and math. With that, I think there's less of an emphasis on the special areas, PE especially, but music and art would be lumped into that as well.

Greatly Needed, but Non-Transparent. As a group, the physical education teachers felt that teacher evaluation in their subject area was very important. Moreover, the teachers felt that teacher evaluation improved accountability and helped with instruction.

Elizabeth is in her first year teaching. She really stressed the fact that accountability was an important component in the teacher evaluation process. She stated,

> I think teachers should be held accountable for teaching what they're supposed to be teaching and for being good teachers and for improving their teaching. On that hand, yes, I think teacher evaluation is important for anyone who's responsible for teaching something to students—that they should be held accountable. In that regard, yes, I think that it's important.

Marie, who has over 10 years' experience, added to the importance that teacher

evaluation had with accountability,

> I feel like if we are gonna be considered part of a child's overall educational
>
> experience, then we as physical educators need to be held accountable for what
>
> we teach. It can't just be, "Oh, today I decide I'm doing this." I think we need to
>
> have a well-rounded curriculum, just like every other subject content area. They
>
> have to cover what they need to cover year after year, and evaluation is one way
>
> for us to be accountable for what we teach.

> The physical education teachers did not understand their current evaluation

system. As a group, they felt that there is no transparency and that they did not

understand what was expected of them. Elizabeth, with one year of experience stated,

> No [laughing]. It's my first year as a PE teacher and it's my first year using it. I
>
> was trained on it a little bit in new teacher orientation. I had a little bit of prior
>
> knowledge about it and then at both my sites there's teachers at the school that are
>
> trained to be mentors for everyone to learn the system. I've gone to different
>
> sessions with them to get a little bit more help and just to understand a little bit
>
> more, but no, I think that the system is kind of exhaustive. There's 60 different
>
> elements you're supposed to know and do and know how to do the right thing you
>
> could be evaluated on. No, I don't feel like [laughing] I completely understand
>
> it.

Brittany (over 10 years) added that,

> It's more applicable to classrooms. Learning how it applies to what we're doing
>
> in PE, like putting up our goals is confusing. Which goal are we using? I think

98

those things are confusing because for us we use our standards. We teach to our

standards every single day.

As a group, the teachers stated that they were uncertain as to what evidence of

student learning and teacher effectiveness was expected from them during the evaluation

process. Johnny (over 10 years) stated,

I think they look at overall behavior management, which I think most of us are

pretty good at. I think in PE, you have to be if you wanna have kids be safe. So

many kids moving in a small environment, behavior management's important. I

think that's one thing that they look at when they come into your classroom. They

can see that right away. Are the kids organized? Are they—are the activities

safe? Are the kids listening? Are they following directions? I think if they're not

in PE, that's gonna stick out right away with all the kids moving, 25 or 30 kids

moving in a small area.

Brittany (over 10 years) also stated,

Well, I think the key areas that they are able to focus on are the management, the

behavior management, seeing how things are set up, making sure it's safe. In the

past, they've done—I've had administrators actually sit and script out things that

I've said that have shown that I'm explaining what I need to explain. I'm giving

positive feedback. I'm giving good feedback to students who might need to make

changes. Things like they'll notice if I go over and I'll—the management piece if

I need to talk to someone quietly or if we need to stop and practice something

over again.

"Who do you Trust".

Physical education teachers as a group did not have confidence in their evaluator. They felt that the administrators do not have the proper training or knowledge of pedagogy in a physical education classroom to give a fair assessment. Sharon, who has eight years' experience stated,

I am not confident. We actually just had a meeting about this with the specials, which is art and music and PE, and she was saying—my evaluator was saying, "If I had to sub for PE, I would just not be confident. I would be worried about safety issues and things like that." I just thought, "If you're not confident teaching my subject, why are coming to evaluate me when you're not?" I'm just not confident in her ability or her ability.

Sheila, with over 10 years of experience stated,

Well, just in talking to a few people, they've presented lessons that maybe take place out in the field with jogging, for instance. They've shared stories of the administrator's literally about 100 yards away with the iPad, 100 yards from where the students are out moving in the field and the PE teacher's moving in the field. They're rating them and I don't know how you could possibly rate someone, or hear what they're saying, or see everything that they're doing when you're 100 yards away from the action. When I think about it in a classroom that would never be, you'd never be 100 yards away from a teacher that you're evaluating.

Jennifer with nine years experience added,

I don't think they know enough about the curriculum or the standards to see,

Oh, you missed standard number two. You're working on this, this, and this." I

don't think they really know those things.

The teachers also agreed that outside expertise and more administrator training is needed

for teacher evaluation specific to physical education to be effective. Kelly (over 10 years)

stated,

I really do think there needs to be training on it. What does this look like in the

classroom? Well, this is what it looks like in PE. We're still doing the same

thing, it just looks a little different. I'm hearing it from the district, I'm hearing it

from my PE teachers who are all talking about this. It needs to cross over. Like I

said, this is what it looks like in the teacher's realm, but this is what we do and

what it looks like here.

Marie (over 10 years) also stated,

I would like to have more of an assessment with someone who knows PE. I

don't know if that would be a supervisor from the University coming over giving

me an evaluation. I would love that, than my principal, who has never been

teaching PE. I'd like someone who has some expertise to help me.

Johnny (over 10 years) stated,

You know, it's hard because I know their time is so—they're pulled in so many

directions. Maybe a special area orientation, maybe before the school year,

maybe a one hour deal where they spend 20 minutes on, "Here's the district's art

curriculum. Here's how most of our teachers in the district teach art. Here's why

they do these certain things." Then another 20 minutes for music and another 20

101

for PE, just so they have a background on what we do, why do it, and just a

general overview of the curriculum that we're all using because it is so much

different than the classroom teachers.

Discussion

This study examined physical education teacher's perceptions of current teacher

evaluation practices in place in their district. The results will be discussed in this section.

Valued, but not Prioritized

This theme is consistent with current literature that suggests physical education

remains a marginalized subject area, and it is not given priority within the school

curriculum (Prince et al., 2008; NASPE, 2012; Puhse & Gerber, 2005; Sheehy, 2011).

Physical education teachers as a group believed that value for physical education was

placed on whether or not the school principal had valued the subject. There is evidence

that administrator's do value physical education and believe that it is very beneficial in

enhancing concentration, decreasing discipline problems, and improving academic

performance (Sallis, McKenzie, Kalody, & Curtis, 1996). Furthermore, the problem may

not lie within the realm of whether or not a school administrator values physical

education, it is more so with the policymakers (Hardman & Marshall, 2000). In an effort

to increase classroom learning time with hopes of higher academic performance, state

and district policymakers have drastically reduced the amount of physical education

students receive in schools (NASPE, 2012). Evidence from the NASPE 2012 shape of the

nation report outline many state level loopholes that reduce the effectiveness of policy

efforts ensuring that quality physical education is present in schools. This evidence may

support a "top down" effect on policies alleviating a lot of the decision making or flexibility school administrators have on physical education.

Greatly Needed, but Non-Transparent

Physical education teachers felt that the current teacher evaluation system was confusing and left them unsure as to what was expected of them. This is similar to what Zimmerman (2003) found in that classroom teachers felt that teacher evaluation systems were not tailored towards them. Moreover, they felt that there was a lack of connection with the teacher and the evaluation process itself (Zimmerman, 2003). Furthermore, relevant to the Performance Improvement Theory guiding this study, physical education teachers are lacking the mental model that is necessary for improvement, therefore leaving them the task of dissecting and interpreting various situations within their current teacher evaluation system (Swanson, 1999).

Physical Educators are NOT Confident in their Evaluators

As a group, the physical education teachers stated that they were not confident that their evaluator could give them a fair and valid assessment. These results are consistent with studies that examined classroom teacher's perceptions of their evaluators. Furthermore, classroom teachers felt that their principals were not adequately qualified to evaluate the subject area. Whence, evaluators who lacked instructional skills were not perceived as having the ability to evaluate instructional content decisions or pedagogical content knowledge (Brandt et al., 2007; Halverson, Kelley, & Kimball, 2004; Prince et al., 2008). This disconnect between teacher and evaluator points towards the increasing evidence that there are many inaccuracies, insufficient training, and lack of support for

administrators expected to execute effective teacher evaluations (Darling-Hammond, Wise, & Kline, 1999; Loup, Garland, Ellet, & Rugutt, 1996; Peterson, 1995).

Strengths and Limitations of the Study

An identifiable strength of this study is that it is one of the first to examine physical education teachers' perceptions of teacher evaluation specific to physical education. This study can serve as a springboard for future research within the area of physical education teacher evaluation. As physical education continues to be held more accountable, and the call for evidence based teacher evaluation persists, there is a need for more research in this area.

There are three identifiable limitations within this study: a) small sample size; b) limited number of interviews; and c) specificity to one school district. This study had a small sample size, and all participants were from one school district. This may have caused a lack of generalizability to the larger population (Locke, Silverman, & Spriduso, 2010). Furthermore, having one interview inhibited the opportunity for follow up questions which may have reduced the chances for richer data (Patel, & Doku, & Tennakoon, 2011).

<div align="center">**Conclusion**</div>

This study examined perceptions of physical education teacher evaluation looking through the lens of the physical educators themselves. It is very apparent that there is a major disconnect with physical educators and current teacher evaluation systems in use. The fact that physical educators do not clearly understand what is expected of them, and do not feel confident in their evaluators are two very concerning issues in today's context of high stakes accountability approaches to teacher evaluation.

Furthermore, as the country continues to rely on physical education as the primary source of physical activity for youth in our schools, a more transparent and consistent evaluation system needs to be in place to ensure proper measures of teacher effectiveness are being performed in physical education. Administrators need to be educated on the ecology of a physical education setting and given more training in regards to content knowledge and pedagogical content knowledge in physical education. Moreover, physical education teachers need to be given extensive training of the evaluation system in use to entail a clear understanding of expectations needed to be met.

Future implications of this study call for researchers, policymakers, and practitioners to work together in order to create an effective physical education teacher evaluation system that is fundamental to quality physical education programs.

CHAPTER 6

SUMMARY

The National Association for Sport and Physical Education (NASPE) describes physical education as an integral part of the total education of every child Kindergarten through 12th grade. The overall goal of a quality physical education program is to produce physically educated persons that will learn to live healthy and active lifestyles (DeJong, Hensley, & Tannehill, 2004). One critical factor in assisting in the quality physical education of our students is having a qualified physical education teacher within the classroom (NASPE, 2007). One critical factor in producing these student learner outcomes is having a qualified physical education teacher who can plan and implement such a program.

With no accountability measures or standardized testing tied to assessment of quality teaching in physical education, formal teacher evaluation systems serve as the primary source of determining teacher performance. This study examined teacher evaluation specific to physical education. Moreover, there were three separate measures aimed to understand physical education teacher evaluation: (a) current practices used by administrators in conducting teacher evaluation on physical education teachers; (b) administrators' perceptions of teacher evaluation specific to physical education; and (c) physical education teacher's perceptions of teacher evaluation.

Teacher evaluation is a standard process used by districts aimed at monitoring teacher's performance. This dissertation study examined physical education teacher evaluation guided by three research questions:

1. What are the current teacher evaluation documents school administrators' use when conducting formal evaluation of physical education teachers?

2. How do the administrators value the evaluation process specific to physical education?

3. How do the physical education teachers view the value of the evaluation process?

Document Analysis

Four commonly used teacher evaluation systems were examined within this study. They included: a) teacher advancement program (TAP) (National Institute for Excellence in Teaching, 2013); b) rewarding excellence in instruction and leadership (REIL) (Maricopa County Education Service Agency, 2013); c) the framework for teaching (FFT) (Danielson, 1996, 2007); and d) the Marzano teacher evaluation model (Marzano Research Laboratory, 2013; Marzano, 2003). All four of these systems target important areas related to quality teaching, and all four emphasize student achievement.

The National Association for Sport and Physical Education (NASPE, 2007) created a teacher evaluation tool to identify the knowledge, skills, and behaviors needed to provide sound instruction in the k-12 physical education classroom. The tool is used as a resource for evaluating teacher behaviors and effectiveness in the physical education classroom. The NASPE tool consists of 5 domains; (1) Instruction; (2) Evidence of Student Learning; (3) Management/Organization; (4) Learning Climate; and (5) Professionalism. Within each domain, there are multiple elements that reflect different teaching behaviors. In total there are 67 elements within the 5 domains.

The NASPE physical education teacher evaluation instrument was used within this study as a basis for resource to determine whether or not the knowledge, skills and

behaviors preferred within a physical education classroom were present within the four teacher evaluation systems reviewed.

There was evidence that a high percentage of key items from the NASPE teacher evaluation instrument were present within the four evaluation systems in question. The two systems with the most connection to the NASPE tool were the TAP and FFT. Moreover, content specific language is missing from the four teacher evaluation systems of inquiry.

Future implications may be the development of an instrument that could compliment currently used systems, or stand alone to measure teacher effectiveness specific to physical education.

Administrators Perceptions

With the absence of standardized testing in Physical education classrooms (NASPE, 2012), the sole measure of teacher performance relies on the teacher evaluation process in place at the district level, as well as the expertise of the school-level administrator charged with conducting the evaluation. Although physical education teachers and the subject itself remain marginalized within the school setting (Sheehy, 2011), teacher evaluation systems remain an important measure in determining effective teaching, professional growth and assisting in student learner outcomes. With school administrators as the likely school site evaluators in physical education teacher evaluation, identifying their perceptions of the overall evaluation process can be important in determining principals' value orientation towards the subject itself and their qualifications for conducting formal high-stakes based teacher evaluations.

Therefore, the purpose of part two of this study was to determine K-8 school administrators' perceptions of conducting formal teacher evaluation of physical education teachers in today's context of high stakes accountability approaches to teacher evaluation. The research question was; how does the administrator perceive the evaluation process specific to physical education? The emphasis of this question was based around four specific foci; (a) perceptions of the value of physical education; (b) perceptions and understanding of measures used to determine teacher effectiveness in physical education; (c) perceptions and understanding of measures used to determine student achievement in physical education; and (d) perceived ability to conduct teacher evaluation in physical education.

An Informal survey and a formal semi-structured interview were used to examine administrators' perceptions of teacher evaluation specific to physical education.

Four common themes emerged from the data; (a) administrators value physical education, but practice within their reality..It is not really on their radar; (b) administrators believe that "good teaching is good teaching"; (c) administrators understand their limitations, and have a desire for improvement of the process; and (d) evaluator training, whats that?

This study is one of the first to examine administrator's perceptions of teacher evaluation specific to physical education. Outcomes from this study suggest more training is needed for administrators in physical education teacher evaluation, and the need for a content specific evaluation instrument. More research will be necessary in the area of physical education teacher evaluation as the call for evidence based outcomes in teacher effectiveness and student achievement are brought to the table.

Teachers Perceptions

        Similar to music and art, physical education remains a marginalized subject receiving low priority and concern within school curriculum (Prince, Schuermann, Guthrie, Witham, Milanowski, & Thorn, 2008). With the obesity epidemic plaguing our nation and school physical education playing a key role in counteracting this dilemma (Pate et al., 2006) , it is more important now than ever before that physical education teachers are receiving quality evaluations and given opportunity for professional growth and development.

        There are no known studies that examine how physical educators perceive current teacher evaluation systems. Therefore, the purpose of this study was to develop a better understanding of current physical education teachers' perceptions of teacher evaluation systems. The research question for this study was; how does the physical education teacher perceive teacher evaluation? The emphasis of this question is based around five specific foci: (a) perceptions of the value of physical education within the school; (b) understanding of quality and standards based physical education; (c) perceptions of teacher evaluation in physical education; (d) confidence of the evaluator; (e) evidence of student learning.

        Informal surveys and formal semi-structured interviews were used to examine teachers' perceptions of teacher evaluation specific to physical education.

        After an in depth analysis, the following three themes emerged from the data: (a) physical education is "valued, but not prioritized; (b) teacher evaluation in physical education is "greatly needed, yet not transparent"; (c) "physical educators are not confident in their evaluators".

This study examined perceptions of physical education teacher evaluation looking through the lens of the physical educators themselves. It is very apparent that there is a major disconnect with physical educators and current teacher evaluation systems in use. Furthermore, as the country continues to rely on physical education as the primary source of physical activity for youth in our schools, a more transparent and consistent evaluation system needs to be in place to ensure proper measures of teacher effectiveness are being performed. Future implications of this study call for researchers, policymakers, and practitioners to work together in order to create an effective physical education teacher evaluation system that is fundamental to quality physical education programs.

REFERENCES

Agam, K., Reifsneider, D., & Wardell, D. (2006). *The Teacher Advancement Program (TAP): National Teacher Attitudes.* Teacher Advancement Program Foundation. Retrieved from www.tapsystem.org/publications/publications

Agam, K., & Wardell, D. (2007). *Annual TAP principal survey: Select findings.* Santa Monica, CA, and Washington, DC: National Institute for Excellence in Teaching. Retrieved from www.tapsystem.org/publications:/publications

Allen, L., Glickman, C., & Hensley, F. (1998). *A search for accountability: The league of professional schools.* Paper presented at the annual meeting of the American Education Research Association, San Diego, CA.

Amrein-Beardsley, A., & Collins, C. (2012). The SAS education value-added assessment system (SAS® EVAAS®) in the Houston Independent School District (HISD): Intended and unintended consequences. *Education Policy Analysis Archives, 20* (12). Retrieved from http://epaa.asu.edu/ojs/article/view/1096

Argyris, C. (1970). *Intervention theory and method*. Reading MA: Addison-Wesley.

Atkins, A. (1996). *Teachers' opinions of the teacher evaluation process.* Unpublished Manuscript. (ERIC Document Reproduction Service No. ED398628

Barosso, C., McCullum-Gomez, C., Hoelscher, D., Kelder, S, & Murray, N. (2005). Self-reported barriers to quality physical education by physical education specialist in Texas. *Journal of School Health,* 75, 313-319.

Berliner, D. (2014). Exogenous variables and value-added assessments: a fatal flaw. *Teachers College Record*. 116(1). Retrieved from http://www.tcrecord.org ID Number: 17293

Bill & Malinda Gates Foundation (2013, January). *Ensuring fair and reliable measures of effective teaching.* Retrieved from Maricopa County Education Service Agency website: http://mcesa.schoolwires.net/domain/127

Blase, J., & Blase, J. (1998). *Handbook of instructional leadership: How really good principals promote teaching and learning*. Thousand Oaks, CA: Corwin Press

Blase, J., & Kirby, P.C. (2000). *Bringing out the best in teachers*: *What effective principals do*. Thousand Oaks, CA: Corwin Press

Bogdan, R.C., & Biklen, S.K. (2007). *Qualitative research for education: An introduction to theory and methods.* Boston: Pearson Education, Inc.

Bolman, L. G., & T. E. Deal. (2003). *Reframing organizations.* San Francisco: Jossey-Bass.

Boyd, R. (1989). Improving teacher evaluations. *Practical Assessment, Research & Evaluation*, *1*(7), 1-3. Retrieved from http://pareonline.net/getvn.asp?v=1&n=7

Brandt, C., Mathers, C., Olivia, M., Brown-Sims, M., & Hess, J. (2007). *Examining district guidance to schools on teacher evaluation policies in the Midwest Region* (REL 2007-No. 030). Retrieved from Institute of Education Sciences website: http://ies.ed.gov/ncee/edlabs/regions/midwest.pdf/rel_2007030.pdf

Braun, H., Chudowsky, N., & Koenig, J. (2010). *Getting value out of value-added*. The National Academies Press. Retreived from http://rlhoover.people.ysu.edu/ClassConnections/core-readings/Getting%20Value%20Out%20of%20Value-Added.pdf

Calabrese, R., Sherwood, K., Fast, J. & Womack, C. (2004). Teachers and principals' perceptions of the summative evaluation conference: An examination of model I theories-in-use. *The International Journal of Educational Management,* 18, 109-117.

Centers for Disease Control and Prevention. (2001). Increasing physical activity: A report on recommendations of the Task Force on Community Preventive Services. *Morbidity and Mortality Weekly Report,* 2001;50(No. RR-18). 1-14.

Centers for Disease Control and Prevention. (2011a). How much physical activity do children need? Available at http://www.cdc.gov/physicalactivity/everyone/guidelines/children.html

Centers for Disease Control and Prevention. (2011b). Youth risk behavior surveillance-united states. Retrieved from www.cdc.gov/HealthyYouth/yrbs/pdf/yrbss10_mmwr.pdf

Centers for Disease Control and Prevention. (2012). Childhood overweight and obesity. Retreived from http://www.cdc.gov/obesity/childhood/

Chait, R. & Miller, R. (2010). *Treating different teachers differently*. Center for American Progress. Retrieved from Maricopa County Education Service Agency: http://mcesa.schoolwires.net/cms/lib03/AZ00001882/Centricity/Domain/127/treatingdifferentteachersdifferently.pdf

Coe, D., Pivarnik, J., Womack, C., Reeves, M., & Malina, R. (2006). Effect of physical education and activity level on academic achievments in children. *Medicine in Sports and Exercise* (pp.1515-1519). DOI:10.1249/01.mss.0000227537.13175.1b

Conger, J., & Kanungo, R. (1994). Charismatic leadership in organizations: Perceived behavioral attributes and their measurement. *Journal of Organizational Behaviors*,15, 439-452.

Danielson, C. (1996). *Enhancing professional practice: A framework for teaching*. Alexandria, VA: ASCD.

Danielson, C. (2007). *Enhancing professional practice: A framework for teaching.* Alexandria. VA: Association for Supervision and Curriculum Development.

Danielson, C. (2012). Teacher evaluation: Whats fair? Whats effective? *Educational Leadership,* 70(3), 32-37.

Danielson, C., & McGreal, T.L. (2000). *Teacher evaluation to enhance professional practice*. Alexandria, VA: Association for Supervision and Curriculum Development & Educational Testing Service.

Darling-Hammond, L., Wise, A., & Klein, S. (1999). *A license to teach: Raising standards for teaching.* San Francisco: Jossey-Bass.

Darling-Hammond, L. (2000). Teacher quality and student achievement: A review of state policy evidence. *Education Policy Analysis Archives*, *8*, 1–44.

Davis, J. W. (1988). The relationship between selected performance evaluation procedures and principals' perceptions about performance evaluation. *Dissertations Abstracts International*, 50/04, 837-A

Dedoose Version (4.5). Web application for managing, analyzing, and presenting qualitative and mixed method research data (2013). Los Angeles, CA: SocioCultural Research Consultants, LLC (www.dedoose.com).

DeJong, G., Hensley, L., & Tannehill, D. (2004). *National standards for physical education*. (2nd ed.). Reston,VA: McGraw-Hill.

Donaldson, M. L. (2009). *So Long, Lake Wobegon? Using Teacher Evaluation to Raise Teacher Quality*. Retrieved from Center for American Progress website: http://www.americanprogress.org/issues/2009/06/pdf/teacher_evaluation.pdf

Glazerman, S., & Seifullah, A. (2010). *An evaluation of the teacher advancement program (TAP) in Chicago.* MATHEMATICA Policy Research Inc. Washington D.C. Retrieved from http://files.eric.ed.gov/fulltext/ED510712.pdf.

Goldrick, L. (2002). *Improving teacher evaluation to improve teaching quality*. Issue Brief. Santa Cruz, CA: Education Policy Studies Division, National Governors Association Center for Best Practices.

Guion, L., Diehl, D., & McDonald, D. (2011). *Triangulation: Establishing the validity of qualitative studies* (FCS6014). Retrieved from Institute of Food and Agricultural Sciences, University of Florida website: http://edis.ifas.ufl.edu/pdffiles/FY/FY39400.pdf

Halverson, R., Kelley, C., & Kimball, S. (2004). Implementing Teacher Evaluation Systems: How Principals Make Sense of Complex Artifacts to Shape Local Instructional Practice. In *Research and Theory in Educational Administration* (pp. 153-188). Retrieved from dm.education.wisc.edu/rrhalverson/intellcont/%20halversonkelleykimball%20110707-1.pdf

Hardman, K., & Marshall, J. (2000). The state and status of physical education in schools in international context. European Physical Education Review, 6, 203-229. DOI: 10.1177/1356336X000063001

Haystead, M. (2010). *Report on professional development: A summary of findings for ten teachers on three critical commitments to school reform*. Englewood, CO: Marzano Research Laboratory**.**

Heneman, H., Milanowski, A., Kimball, S., & Odden, A. (2006). *Standards-based teacher evaluation as a foundation for knowledge- and skill-based pay*. Philadelphia, PA: Consortium for Policy Research in Education.

Herlihy, C. (October, 2012). *State and local efforts to investigate the validity and reliability of scores from teacher evaluation systems.* Unpublished paper presented at the high stakes teacher evaluation: high cost-big losses conference Tucson, Arizona. Retrieved from http://www.gse.harvard.edu/cepr-resources/files/news-events/ncte-valid-reliable-high-stakes-conference-herlihy-hill.pdf

Hill, C., Kapitula, L., & Umland, K. (2011). A validity argument approach to evaluating teacher value-added scores. *American Educational Research Journal*, 48(3), 794, 794-831.

Hill, H., & Herlihy, C. (2011). Prioritizing teaching quality in a new system of teacher evaluation. *American Enterprise Institute for Public Policy Research*, *9* (November), Retrieved from www.aei.org

Hinchey, P. (2010). *Getting Teacher Assessment Right:What Policymakers Can Learn From Research*. Retrieved from National Education Policy Center website: http://nepc.colorado.edu/publication/getting-teacher-assessment-right.

Huffman, D., & Thomas, K. (2003). Relationship between professional development, teachers' instructional practices, and the achievement of students in science and mathematics. *School Science and Mathematics*, 103, 378-387.

Hussey, J., & Khandaker, N. (2012). *Measuring and Increasing Educator Effectiveness: A Balanced Approach*. Retrieved from http://Battelleforkids.org

Jacob, B., & Lefgremn, L. (2008). Can principals identify effective teachers? Evidence on subjective performance evaluation in education. *Journal of Labor Economics, 26*(1), 101-136.

Jerald, C. (2009). *Aligned by design: How teacher compensation reform can support and reinforce other educational reforms*. Washington, DC: Center for American Progress.

Kane, T., Taylor, E., Tyler, J., & Wooten, A. (2010). *Identifying effective classroom practices using student achievement data (paper number 15803).* National Bureau of Economic Research. Cambridge, MA.

Kennedy, M.(2008). "Sorting Out Teacher Quality." *Phi Delta Kappan, 90*(1), 59-63.

Kersten, T., & Israel, M. (2005). Teacher evaluation: Principals' insights and suggestions for improvement. *Planning and Changing*, 36(1 & 2), 47-67.

Kimball, S. (2002). Analysis of feedback, enabling conditions and fairness perceptions of teachers in three school districts with new standards-based evaluation systems. *Journal of Personnel Evaluation in Education*, 16, 241-268.

Kimball, S., & Milanowski, A. (2009). Examining teacher evaluation validity and leadership decision making within a standards-based evaluation system. *Educational Administration Quarterly, 45(1), 34-70.*

Koppich, J. (2008). *Reshaping teacher policies to improve student achievement*. Berkeley, CA: Policy Analysis for California Education.

Kvale, S., & Brinkman, S. (2009). *Interviews, learning the craft of qualitative research interviewing.* Thousand Oaks, CA: Sage Publications.

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

116

Levi, J., Segal, I., & Juliano, C. (2006). *F as in fat:How obesity threatens america's future.* Washington, DC: Trust for Americas Health.

Levine, A. (2005). *Educating school leaders*. Washington DC: The Education School Project. Retreived from http://files.eric.ed.gov/fulltext/ED504144.pdf

Lincoln, Y., & Guba, E. (1985). *Naturalistic inquiry*. Beverly Hills, CA: Sage

Little, O. (2009). *Teacher evaluation systems: The window for opportunity and reform*. Washington D.C.: NEA Research.

Little, O., Goe, L., & Bell, C. (2009). *A practical guide to evaluating teacher effectiveness.*National Comprehensive Center for Teacher Quality: Washington DC. Retreived from http://files.eric.ed.gov/fulltext/ED543776.pdf

Locke, L. (1974). *The ecology of the gymnasium: What the tourists never see.* Paper presented at the meeting for the Southern Association for Physical Education for College Women, Gatlinburg, TN.

Locke, L.F., Silverman, S.J., & Spirduso, W.W. (2010). *Reading and understanding Research* (3rd ed). Thousand Oaks, CA: Sage Publications.

Lounsbery, M., McKenzie, T., Trost, S., & Smith, N. (2011). Facilitators and barriers to adopting evidence based physical education in elementary schools. *Journal of Physical Activity and Health*, 8(suppl 1), S17-S25.

Loup, K. S., Garland, J. S., Ellett, C. D., & Rugutt, J. K. (1996). Ten years later: Findings from a replication of a study of teacher evaluation practices in our 100 largest districts. *Journal of Personnel Evaluation in Education*, *10*(3), 203–226.

March, J., & Simon, H. (1993). *Organizations.* (2nd ed.). Cambridge, MA: Blackwell Business.

Maricopa County Education Service Agency (MCESA). (2013). *Rewarding excellence in instruction and leadership*. Retrieved from http://mcesa.schoolwires.net/domain/106

Marion, S., & Buckley, K. (2011). *Approaches And Considerations For Incorporating Student Performance Results From "Non-Tested" Grades And Subjects Into Educator Effectiveness Determinations*. Cambridge MA: National Center for the Improvement of Educational Assessment.

Marzano, R., Pickering, D., & Pollock, J. (2001). *Classroom Instruction that Works*. Alexandria, VA: Association for Supervision and Curriculum Development.

Marzano, R. (2003). *What Works in Schools*. Alexandria, VA: Association for Supervision and Curriculum Development.

Marzano, R., Pickering, D., & Marzano, J. (2003). *Classroom management that works.* Alexandria, VA: Association for Supervision and Curriculum Development.

Marzano, R. (2006). *Classroom Assessment and Grading that Works*. Alexandria, VA: Association for Supervision and Curriculum Development.

Marzano, R. (2007). *The Art and Science of Teaching*. Alexandria, VA: Association for Supervision and Curriculum Development.

Marzano Research Laboratory. (2010). *Phase II – What works in Oklahoma schools? Retrieved* from http://www.marzanoresearch.com/documents/Phase_II_OK_State_Report.pdf

Marzano, R., & Haystead, M. (2010). *Final report: A second year evaluation study of Promethean ActivClassroom.* Englewood, CO: Marzano Research Laboratory. Retrieved from http://www.marzanoresearch.com/documents/Continuation_Study_2010.pdf

Marzano,R., & Haystead, M. (2011). *2010-2011 Adams 50 instructional model study.* Englewood, CO: Marzano Research Laboratory. Retrieved from http://www.sbsadams50.org/wp-content/uploads/2011/10/ACSD50-Instructional-ModelFinal-Validity-Report-2011.pdf

Marzano Research Laboratory. (2011*). Research base and validation studies on the marzano evaluation model.* Retrieved from http://www.marzanoevaluation.com/files/Research_Base_and_Validation_Studies_Marzano_Evaluation_Model.pdf

Marzano,R., & Haystead, M. (2011). *2010-2011 Adams 50 instructional model study.* Englewood, CO: Marzano Research Laboratory. Retrieved from http://www.sbsadams50.org/wp-content/uploads/2011/10/ACSD50-Instructional-ModelFinal-Validity-Report-2011.pdf

Marzano, R., Frontier, T., & Livingston, D. (2011). *Effective Supervision: Supporting the art and science of teaching*. Alexandria, VA: Association for Supervision and Curriculum Development.

Marzano, R., Toth, M., & Schooling, P. (2012). *Examining the role of teacher evaluation in student achievement: Contemporary research base for the Marzano Causal Teacher Evaluation Model.* Retrieved from http://www.marzanocenter.com/files/MC_White_Paper_20120424.pdf

118

Marzano Research Labortory. (2013). *Marzano Teacher Evaluation Model.* Retrieved from www.marzanoevaluation.com

McKenzie, T., Marshall, S., Sallis, J., & Conway, T. (2000). Student activity levels, lesson context, and teacher behaviors during middle school physical education. *Research Quarterly,* 71, 249-259.

Mercier, K., & Doolittle, S. (2013). Assessing student achievement in physical education for teacher evaluation. *Journal of Physical Education, Recreation & Dance,* 84(3), 38-42, DOI: 10.1080/07303084.2013.767721

Millman, J. (1981). *Handbook of teacher evaluation.* Beverly Hills, CA: Sage Publications.

National Association For Sport And Physical Education (NASPE). (2007). *Physical Education Teacher Evaluation Tool.* Reston, VA: Author

National Association for Sport and Physical Education (NASPE). (2010). *Shape of the nation report: Status of Physical Education in the USA.* Reston, VA: American Alliance for Health, Physical Education, Recreation and Dance.

National Association for Sport and Physical Education (NASPE)(2012a). *Why children need physical education.* Retrieved from http://www.aahperd.org/naspe/publications/teachingTools/whyPE.cfm

NASPE & American Heart Association. (2012). *Shape of the nation report: Status of Physical Education in the USA.* Reston, VA: American Alliance for Health, Physical Education, Recreation and Dance.

National Commission on Excellence in Education. (1983). *A nation at risk: The imperative for educational reform.* Washington D.C.: United States Department of Education.

National Council on Teacher Quality (2011). *State of the States:Trends and Early Lessons on Teacher Evaluation and Effectiveness Policies* (October 2011). Retrieved from http://www.nctq.org

National Education Association. (2014). New policy statement on teacher evaluation and accountability-Adopted as amended. Retrieved from http://www.nea.org/grants/46326.htm

National Institute for Excellence in Teaching. (2013). *Teacher Advancement Program.* Retrieved from www.tapsystem.org.

Olivia, M., Mathers, C., & Laine, S. (2009). Effective evaluation. *Principle Leadership*. Retrieved from http://www.nassp.org/Portals/0/Content/59182.pdf

Ovando, M. (2001). Teachers' perceptions of a learner-centered teacher evaluation system. *The Journal of Personnel Evaluation in Education.* 15, 213-231.

Painter, S. (2001).Barriers to evaluation: Beliefs of elementary and middle school principals. *Planning and Changing*, 32(1), 58-70.

Pate, R., Davis, M., Robinson, T., Stone, E., McKenzie, T., & Young, J. (2006). Promoting physical activity in children and youth: A leadership role for schools: A scientific statement from the American Heart Association Council on Nutrition, Physical Activity, and Metabolism (Physical Activity Committee) in collaboration with the Councils on Cardiovascular Disease in the Young and Cardiovascular Nursing. *Circulation,* 114(11), 1214-1224.

Patel, M., Doku, V., & Tennakoon, L. (2003). Challenges in recruitment of research participants. *Advances in Psychiatric Treatment*, 9.

Peterson, A. (1995). Teacher-researcher compatibility: A view from both sides. *Remedial and Special Education.* 16(6)*,* 364-367.

Peterson, K. (2000). *Teacher evaluation: A comprehensive guide to new directions and practice* (2nd ed). Thousand Oaks, CA: Corwin Press.

Peterson, K. (2001). *The professional development of principals: Innovations and opportunities.* Paper commissioned for the first meeting of the National Commission for the Advancement of Educational Leadership Preparations, Racine, WI.

Prince, C., Schuermann, P., Guthrie, J., Witham, P., Milanowski, A., & Thorn, C. (2008). *The other 69 percent: Fairly rewarding the performance of teachers of non-tested subjects and grades*. Washington, DC: Center for Educator Compensation Reform, U.S. Department of Education.

Protheroe, N. (2002). Improving instruction through teacher observation. *Principal*, 82(1), 48-51.

Puhse, U., & Gerber, M. (2005). *International comparison of physical education: concepts, problems, prospects.* Oxford, UK: Meyer & Meyer Sport.

Rothstein, R., Ladd, H., Ravitch, D., Baker, E., Barton, P., Darling-Hammond, L., Haertel, E., Linn, R., Shavelson, R., & Shepard, L. (2010). *Problems with the use of student test scores to evaluate teachers | Economic Policy Institute*. Retrieved from http://www.epi.org/publication/bp278

Sallis, J., & McKenzie, T. (1991). Physical educations role in public health. *Research Quarterly for Exercise and sport*, 62, 124-137.

Sallis, J., McKenzie, T., Kolody, B., & Curtis, P. (1996). Assessing district administrators' perceptions of elementary school physical education. *Journal of Physical Education, Recreation, and Dance,* 67(8), 25-29.

Sallis, J., McKenzie, T., Beets, M., Beighle, A., Erwin, H., & Lee, S. (2012). Physical educations role in public health: steps forward and backward over 20 years and hope for the future. *Research Quarterly for Exercise and Sport*, 83, 125-135.

San Diego State University. (2007). *Physical education matters*: *A report from the california endowment*. Los Angeles CA. Retrieved from http://www.calendow.org/uploadedfiles/publications/by_topic/disparities/obesity_ and_diabetes/pe%20matters%20long%20versionfinal.pdf

Sartain, L., Stoelinga, S., & Brown, E. (2009). *Evaluation of the Excellence in Teaching Pilot: Year One Report.*Chicago, Ill.: Consortium on Chicago School Research.

Sartain, L., Stoelinga, S., & Krone, E. (2010) *Rethinking teacher-evaluation: Findings from the first year of the Excellence in Teaching Project in Chicago Public Schools.* Chicago: Consortium on Chicago School Research at the University of Chicago Urban Education Institute.

Schacter, J., & Thum, Y. (2004). Paying for high-and low-quality teaching. *Economics of Education Review,* 23(4), 411-430.

Scriven, M. (l981). Summative teacher evaluation. In J. Millman (ed.). *Handbook of Teacher Evaluation* (pp. 244).Beverly Hills: Sage Publications.

Shakman, K., Riordan, J., Sanchez, M., DeMeo-Cook, K., Fournier, R., & Brett, J. (2012). *An examination of performance-based teacher evaluation systems in five states.* (Issues & Answers Report, REL 2012-No. 129). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Northeast and Islands. Retrieved from http://ies.ed.gov/ncee/edlabs/regions/northeast/pdf/rel_2012129.pdf

Sheehy, D. (2011): Addressing parents' perceptions in the marginalization of Physical Education. *Journal of Physical Education, Recreation & Dance,* 82(7), 42-56.

Sheppard, B. (1996). Exploring the transformation nature of instructional leadership. *Alberta Journal of Educational Research*, 42(4), 325-344.

Solomon, L., White, J., Cohen, D., & Woo, D. (2007). *The effectiveness of the teacher advancement program*. Santa Monica, CA, and Washington D.C: National Institute for Excellence in Teaching. Retrieved 04/02/13 from www.tapsystem. Org/publications/publications. Taf?page=reports.

Starbuck, W., & Milliken, F. (1988). Executive perceptual filters: What they notice and how they make sense. In D. Hambrick (ed.). *The executive effect: Concepts and methods for studying top managers:* (pp.35-65). Greenwich, CT: JAI Press.

Stronge, J., & Tucker, P. (2003). *Handbook on teacher evaluation: Assessing and improving performance*. Larchmont, NY: Eye On Education.

Swanson, R. (1999). The foundations of performance improvement and implications for practice. In R.Torraco (Ed.), *The theory and practice of performance improvement:* (pp. 1-25). doi:10.1177/152342239900100102.

Treslan, D., & Ryan, J. (1986). Perceptions of principals' authority bases. *Morning Watch*, 14, 1-2, 28-42.

Trost, S. (2009). Active education: Physical education, physical activity and academic performance. *Active Living Research*, Summer 2009. Retrieved from www.activelivingresearch.com/alr/alr/files/Active_Ed.pdf

Tziner, A., Murphy, K., & Cleveland, J. (2001). Relationship between attitudes toward organizations and performance appraisal systems and rating behavior. *International Journal of Selection and Assessment, 9*, 226-239.

U.S. Department of Education (2002). No Child Left Behind (NCLB) Act of 2001(2002).-Pub. L. No. 107-110, 115 §1425.

U.S. Department of Education. (2009). *Race to the top program executive summary*. Retrieved from website: www.2.ed.gov/programs/racetothetop/executive-summary.pdf

U.S Department of Health and Human Services. (2008). *Physical activity guidelines for americans*.Retreived from http://www.health.gov/PAGuidelines/pdf/pagude.pdf

Valentine, J. (1992). *Principles and practices for effective teacher evaluation.* Needham Heights, MA: Allyn and Bacon.

Wang, C., & Ha, A. (2008). The teacher development in physical education: A review of the literature. *Asian Social Science,* 4(12), 3-18.

Williams, L., & Rink, R. (2003). Teacher competency using observational scoring rubrics. *Journal of Teaching in Physical Education,* 22, 552-572.

Zimmerman, S. (2003). Five steps for improving teacher evaluation: Focusing on the continual improvement of teaching and learning. *The Professional Educator,* 25(2), 43-54.

Zimmerman, S., & Deckert-Pelton, M. (2003). Evaluating the evaluators: Teachers' perceptions of the principals' role in professional evaluation. *NASSP Bulletin, 87(636),* 28-37.

APPENDIX A

RECRUITMENT LETTER FOR SUPERINTENDENTS

July 24, 2012

Dear Superintendent:

My name is Hans van der Mars; I am a professor in Arizona State University's Mary Lou Fulton Teachers College. One of my doctoral students and I are conducting a research project aimed at assessing the structure and content of the tools currently used when conducting formal teacher evaluation of the physical education teachers in your district. This email is a formal request for the contact information of the individual responsible for teacher evaluation within your district.
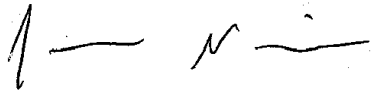
This study has been approved through the Institutional Review Board (IRB) at Arizona State University. The contact information will only be used in the requesting of a copy of the evaluation tool.   Any information received from your district will remain anonymous and kept strictly between our research team.

If you have any questions or concerns regarding this request, please feel free to contact Dr. Hans van der Mars (480-727-1653 or hans.vandermars@asu.edu) or Jason Norris
(253-576-7987 or Jason.norris@asu.edu).   We thank you for your assistance in helping us complete this project, and appreciate your time and cooperation.

Sincerely,

Hans van der Mars, PhD.
Professor in Physical Education

Jason Norris, MPE.
PhD. Student

APPENDIX B

COMPREHENSIVE TABLE OF TEACHER EVALUATION SYSTEM ALLIGNMENT

Table X

TITLE

Domain	NASPE Key Items	Danielson	TAP	Marzano	REIL
Instruction	Lesson introduction is appropriate		†		
Instruction	Learning expectations/objectives/instructional goals are clearly communicated to students	†	†	†	†
Instruction	Content is accurate and current	†	†		†
Instruction	Content and tasks are developmentally appropriate and properly sequenced	†	†	†	†
Instruction	Content and tasks are presented concisely and clearly, emphasizing key elements	†	†	†	†
Instruction	Engages students in learning by enabling all learners to participate through multiple modalities (opportunities to practice the skill).	†	†	†	†
Instruction	Opportunities for teachable moments are recognized and utilized	†	†	†	†
Instruction	Instruction is differentiated for all learners (Accommodations and modifications are made for students with disabilities or varied learning styles).	†	†		
Instruction	Specific, meaningful and timely feedback is provided to students (e.g., peformance, efforts & positive contributions)	†	†	†	†
Instruction	Content is linked to and promotes the transfer of learning within physical education units and among other subject content areas	†	†	†	†
Instruction	Student performance is continually assessed to guide instruction	†	†		†
Instruction	Independent learning is promoted ,encouraged, and reinforced through daily assessments	†	†		†
Instruction	Lesson pace is appropriate	†	†	†	
Instruction	Appropriate closure is provided		†		
Evidence of Student Learning	Assessment is based on mastery of learning expectations which are aligned with local, state and national standards		†	†	
Evidence	There is ongoing formal and informal	†	†	†	†

127

Category	Indicator				
of Student Learning	assessment				
Evidence of Student Learning	Assessment criteria is communicated to students	†	†	†	
Evidence of Student Learning	Multiple assessment strategies and tools are used (formative and summative) to monitor student learning	†	†		†
Evidence of Student Learning	Students are able to articulate relevance and transfer of learning		†	†	†
Evidence of Student Learning	Student progress is documented in a retrievable record-keeping system	†			
Evidence of Student Learning	Student progress and achievement is communicated regularly to relevant stakeholders	†		†	†
Management/Organization	Lesson plans and curriculum are aligned w/ current local, state, and national standards		†	†	
Management/Organization	Instructional area is safe, orderly, and supports learning activities	†	†	†	
Management/Organization	Adequate and developmentally appropriate equipment is accessible and utilized	†	†		
Management/Organization	Instructional support materials are utilized to enhance the lesson.	†	†	†	
Management/Organization	Students understand and adhere to class rules, routines and behavioral expectations	†	†	†	†
Management/Organization	Class routines maximize instructional time	†	†	†	†
Management/Organization	There is a behavior management plan that is fair, firm, and equitable	†	†	†	†
Management/Organization	Appropriate behaviors are reinforced consistently	†	†	†	†
Management/Organization	Students are actively monitored and closely supervised using effective management strategies	†	†		†
Management/Organization	Students are appropriately grouped	†	†	†	†
Management/Organization	Effective and smooth transitions are apparent	†	†	†	
Management/Organization	Allocated time is used effectively and efficiently allowing students to remain focused on the lesson and task expectations.	†	†		†

Managem ent/Organi zation	Students are engaged in relevant, meaningful physical activity a minimum of 50-60 % of the instructional time.				
Managem ent/Organi zation	Accurate records are maintained	†			†
Learning Climate	Lifelong physical activity and skillful movement are promoted			†	
Learning Climate	There is a safe, secure, learning environment that promotes, success, appropriate risk taking, positive self-expression and enjoyment	†	†		†
Learning Climate	High expectations for learning and behavior are evident	†	†		†
Learning Climate	Climate of courtesy and respect is established	†	†	†	†
Learning Climate	Students support the learning of others	†	†	†	†

Note. † = present in the evaluation system.

APPENDIX C

INFORMED CONSENT

INFORMED CONSENT

Current Practices and Perceptions of Physical Education Teacher Evaluation Systems

<u>**INTRODUCTION**</u>
The purposes of this form are to provide you (as a prospective research study participant) information that may affect your decision as to whether or not to participate in this research and to record the consent of those who agree to be involved in the study.

<u>**RESEARCHERS**</u>
Hans van der Mars, PhD., with the Department of Physical Education at Arizona State University along with Jason Norris, Doctoral Candidate., have invited your participation in a research study.

<u>**STUDY PURPOSE**</u>
Given the current focus on high–stakes accountability in American schools, and the status of physical education within that context, the purpose of this study is to:
a) Gain understanding of current teacher evaluation practices used when evaluating physical education teachers,
b) Reveal perceptions of how Physical Education teachers value the evaluation process, and
c) Reveal perceptions of how school administrators value the evaluation process, specific to physical education.

<u>**DESCRIPTION OF RESEARCH STUDY**</u>
If you decide to participate, then you will join a study involving research of Physical Education teacher evaluations. You will be asked to participate in a one on one interview with the researcher, focused on your perceptions of current teacher evaluation practices, along with other questions aimed at understanding teacher evaluations in your school/district. It will be your discretion whether or not you answer any of the questions asked during the interview. This interview will be audio recorded and transcribed at a later date. Notes will be taken during the interview process.

If you say YES, then your participation will last for the duration of two (2) interviews. The interview will last up to one hour each at your school or alternate location convenient to you.
There are a total of 10 subjects that will participate in this study from your within your district.

<u>**RISKS**</u>
There are no known risks from taking part in this study, but in any research, there is some possibility that you may be subject to risks that have not yet been identified.

BENEFITS
This study will potentially help teacher educators and future teachers to have a better understanding of the potential value and limitations of teacher evaluation practices. This study may also assist in the preparation of future teachers for on the job evaluations. Through this study, administrators may begin to see where they are lacking knowledge and/or awareness related to the evaluation of Physical Education teachers. Outcomes may also help administrators in term of Physical Education teachers' professional development, best Physical Education teacher practices and student learner outcomes.

CONFIDENTIALITY
All information obtained in this study is strictly confidential. The results of this research study may be used in reports, presentations, and publications, but the researchers will not identify you, your school or your district. In order to maintain confidentiality of your records, Hans van der Mars will use a coding system and pseudonyms when participants are identified. All information and all original identifying records will be stored in a locked cabinet at the ASU Polytechnic Campus in Santa Catalina Hall room 330 S and later destroyed upon the completion of data collection and analysis.

WITHDRAWAL PRIVILEGE
It is ok for you to say no. Even if you say yes now, you are free to say no later, and withdraw from the study at any time.

Your participation is voluntary. Nonparticipation or withdrawal from the study will have no negative effect on your relationship with Arizona State University.
COSTS AND PAYMENTS
There is no payment or cost for your participation in the study.

VOLUNTARY CONSENT
Any questions you have concerning the research study or your participation in the study, before or after your consent, will be answered by:

Hans van der Mars, PhD
Professor of Physical Education
Arizona State University-Polytechnic Campus
Department of Physical Education
7271 E. Sonoran Arroyo Mall, Santa Catalina Hall
Room 330S
Mesa, AZ 85212
(480)727-1653
email: hans.vandermars@asu.edu

OR

Jason Norris
Doctoral Candidate

Arizona State University-Polytechnic Campus
Department of Physical Education
7271 E. Sonoran Arroyo Mall, Santa Catalina Hall
Room 350G
Mesa, AZ 85212
(253)576-7987
Email: Jason.norris@asu.edu

If you have questions about your rights as a subject/participant in this research, or if you feel you have been placed at risk; you can contact the Chair of the Human Subjects Institutional Review Board, through the ASU Office of Research Integrity and Assurance, at 480-965 6788.

This form explains the nature, demands, benefits and any risk of the project. By signing this form you agree knowingly to assume any risks involved. Remember, your participation is voluntary. You may choose not to participate or to withdraw your consent and discontinue participation at any time without penalty or loss of benefit. In signing this consent form, you are not waiving any legal claims, rights, or remedies. A copy of this consent form will be given (offered) to you.

Your signature below indicates that you consent to participate in the above study.

_____   _____   _____
Subject's Signature   Printed Name   Date

_____   _____   _____
Legal Authorized Representative   Printed Name   Date
(if applicable)

INVESTIGATOR'S STATEMENT
"I certify that I have explained to the above individual the nature and purpose, the potential benefits and possible risks associated with participation in this research study, have answered any questions that have been raised, and have witnessed the above signature. These elements of Informed Consent conform to the Assurance given by Arizona State University to the Office for Human Research Protections to protect the rights of human subjects. I have provided (offered) the subject/participant a copy of this signed consent document."

Signature of Investigator_____

APPENDIX D

ADMINISTRATOR SURVEY

Consent
0=no; 1=yes

Ethnicity
0=white; 1=Hispanic

Gender
0=female; 1=male

Years as Admin
0=1-5; 1=5-10; 2=over 10
Taught Physical Education
0=no; 1=yes

1-Physical education is just as important to whole child development as are "core subjects" (e.g., Mathematics): 19/20 (95%) agreed or strongly agreed with this statement
1=strongly disagree; 2=disagree; 3=agree; 4=strongly agree

2-Teacher evaluation is as important for physical education teachers as it is for teachers of "core subjects":
1=strongly disagree; 2=disagree; 3=agree; 4=strongly agree

3-Physical education teachers should be held to the same expectations as teachers of other school subjects:
1=strongly disagree; 2=disagree; 3=agree; 4=strongly agree

4-Teacher evaluations are a useful tool for professional growth in physical education:
1=strongly disagree; 2=disagree; 3=agree; 4=strongly agree

5-Teacher evaluations are a reliable measure of teacher effectiveness in physical education:
1=strongly disagree; 2=disagree; 3=agree; 4=strongly agree

6-Current teacher evaluation practices impact teachers' classroom practices positively:
1=strongly disagree; 2=disagree; 3=agree; 4=strongly agree

7-Evidence of student growth and achievement is (or "should be"??) an important factor of teacher evaluation in physical education:
1=strongly disagree; 2=disagree; 3=agree; 4=strongly agree

8-My district's teacher evaluation system (e.g., Marzano) used within my district can accurately assess/determine the pedagogical content knowledge of physical education teachers:
1=strongly disagree; 2=disagree; 3=agree; 4=strongly agree

9-I am skilled in accurately employing the current formal teacher evaluation tool (e.g., Marzano) when evaluating the performance of classroom teachers:
1=strongly disagree; 2=disagree; 3=agree; 4=strongly agree

10-I am skilled in accurately employing the current formal teacher evaluation tool (e.g., Marzano) when evaluating the performance of physical education teachers:
1=strongly disagree; 2=disagree; 3=agree; 4=strongly agree

11-I am highly skilled in using the data collected through the districts's current teacher evaluation system for use in post-observation feedback sessions with my physical educator(s):
1=strongly disagree; 2=disagree; 3=agree; 4=strongly agree

12-Feedback sessions/Post evaluation conferences are valuable to the professional growth of my physical education teacher(s):
1=strongly disagree; 2=disagree; 3=agree; 4=strongly agree

What, if any, improvements can or should be made to make the teacher evaluation system more useful and effective for evaluating your physical educator(s)?

Have you ever had to give a physical education teacher a negative score or put them on a personal improvement plan (PIP)? If so, please explain.

To further investigate physical education teacher evaluation practices, would you be interested in participating in a brief interview? Your participation is important and would be greatly appreciated. The interview would last no longer than one hour.

If you answered Yes, please supply your email address and I will contact you for availability. Thank you.

APPENDIX E

INTERVIEW QUESTIONS FOR ADMINISTRATORS

a) Perceptions of the value of physical education; b) perceptions and understanding of measures used to determine teacher quality in physical education; c) perceptions and understanding of measures used to determine student achievement in physical education; d) perceived ability to conduct teacher evaluation in physical education.

1. How important is Physical Education within the school curriculum? Why or Why not?
    a. Do you feel that it is just as important as the various classroom subjects? Why or why not?
    b. What makes PE just as important?

2. What curricular model(s) is used within your Physical Education department?
    a. What can you tell me about this particular model?

3. What is your understanding of quality physical education?

4. Is formal professional development provided to teachers in Physical Education and if so what types? If not, why is this not occurring?

5. What is the importance of teacher evaluation in physical education?

6. How often are Physical Education teachers supposed to be evaluated?
    a. Do you feel that is a sufficient number? Why or why not?

7. Specific to your districts evaluation tool, what are key elements you look for when scoring the physical education teacher?

8. What evidence should be used in measuring teacher effectiveness in physical education?

9. To what extent is student achievement factored into the teachers' evaluation?

10. What evidence should be used to measure student growth outcomes in physical education?

11. Your district is moving to (or is employing) a value added model of teacher evaluation. How comfortable are you in employing this type of teacher evaluation?
12. What do you see as the positive aspects of VAM's of teacher evaluation?

13. What type of feedback is the focus of the post evaluation/feedback session?
    a. Do you feel that these sessions are beneficial to the growth of the teacher? Why or why not?

14. What rewards do teachers receive for a positive evaluation?

15. What are some consequences teachers receive for a negative evaluation? What might a typical improvement plan look like for the teacher in question?

16. Some experts have argued that using generic teacher evaluation tools such as Danielson or Marzano) lack the subject matter specificity to be sensitive to teaching skills that reflect pedagogical content knowledge of teachers. Do you agree or disagree? Explain.

17. How would you describe the quality of training you received to do teacher evaluation during your school administrator certification training? What was good about it? What were the shortcomings?

18. What do you see as your strengths as a school administrator when it comes to mentoring/evaluation of physical educators?

19. How confident are you that you can offer a fair and valid evaluation of a physical education teacher given the unique context of the subject matter being taught?

APPENDIX F

PHYSICAL EDUCATION TEACHER SURVEY

Consent
0=no ;1=yes

Ethnicity(23 caucasian; 1 other, Biracial; 2 no answer)
0=caucasian; 1=other,biracial)

Gender (11Male;13 female; 2 no answer)
0=female; 1=male

Years as Teacher
1=1-5; 2=5-10;3=Over 10

1-Physical education is a priority in my school curriculum:
1=strongly disagree; 2=disagree; 3=agree; 4=strongly agree

2-Teacher evaluation is necessary in physical education:
1=strongly disagree; 2=disagree; 3=agree; 4=strongly agree

3-Physical Education teachers need to be held to the same expectations as teachers of
other school subjects:
1=strongly disagree; 2=disagree; 3=agree; 4=strongly agree

4-Teacher evaluations are a useful tool for professional growth in physical education:
1=strongly disagree; 2=disagree; 3=agree; 4=strongly agree

5-Teacher evaluations are a reliable measure of teacher effectiveness in physical
education:
1=strongly disagree; 2=disagree; 3=agree; 4=strongly agree
10-Teacher evaluation does have an effect on teacher practices:
1=strongly disagree; 2=disagree; 3=agree; 4=strongly agree

6-Evidence of student growth and achievement is an important factor of teacher
evaluation in physical education:
1=strongly disagree; 2=disagree; 3=agree; 4=strongly agree

7-I completely understand the current teacher evaluation system used within my district
and know exactly what is expected of me:
1=strongly disagree; 2=disagree; 3=agree; 4=strongly agree

8-I am confident that my administrator/evaluator is able to determine my effectiveness as
a physical educator:
1=strongly disagree; 2=disagree; 3=agree; 4=strongly agree

9-Feedback sessions/Post evaluation conferences are valuable to my growth as a physical
education teacher:

1=strongly disagree; 2=disagree; 3=agree; 4=strongly agree

10-What are some improvements, if any, that would be of importance if used within the current physical education teacher evaluation systems?

Have you ever received a negative score or have been put on a personal improvement plan (PIP)? If so, please explain.

To further investigate physical education teacher evaluation systems, would you be interested in participating in a brief interview? The interview would last no longer than one hour.

If you answered yes, please supply your email address and I will contact you for availability. Thank you.

APPENDIX G

INTERVIEW QUESTIONS FOR PHYSICAL EDUCATION TEACHERS

a) Perceptions of the value of physical education within the school; b) understanding of quality and standards based physical education; c) perceptions of teacher evaluation in physical education; d) confidence of the evaluator; e) evidence of student learning

1. How important is Physical Education within the school curriculum? Why or Why not?
   a. Where do you feel physical education falls priority wise within your school (e.g., is physical education as important as "core subjects")

2. Do you feel that standards - based Physical Education is important? Why or why not?
   a. How do you know that your students are meeting the state content standards?

3. Do you follow a particular curricular model when teaching physical education in your classroom? Who decided on that particular curriculum and why?
   a. How comfortable are you teaching this curricular model?

4. Do you feel that teacher evaluation systems are necessary in Physical Education? Why or why not?

5. How do you prepare for your evaluation by the administrator?
   a. Do you feel a sense of anxiousness or nervousness prior to your observations? Why or why not?

6. What are some key areas you feel are targeted by your administrator during your observations?

7. What are some instructional/learning outcome areas you feel are necessary towards receiving a positive score?

8. Do you completely understand the teacher evaluation system (e.g., Marzano) that is used within your district? Explain

9. Given the unique context of the subject matter being taught, how confident are you in your administrator that (s) he is able to determine your effectiveness as a physical educator? Please explain.
   a. What changes if any

10. Do you feel that the feedback session/Post evaluation meetings are valuable to your growth as a teacher? Explain why.

11. Have you ever received a negative evaluation score or put on a PIP? Explain

12. The call for using "evidence-based" physical education curricula is more common today. Do you believe that the curriculum in place in the district today allows you to demonstrate that students in your program learn something worthwhile?

13. The new approach to teacher evaluation in the district requires physical educators to demonstrate that their students have learned. A) What are your feelings about this requirement? B) To what extent are you prepared to fulfill this requirement?

14. What are some improvements, if any, that would be of importance if used within the current Physical Education teacher evaluation systems?