Copyright

by

Vanessa Leigh Errisuriz

2016

The Dissertation Committee for Vanessa Leigh Errisuriz Certifies that this is the approved version of the following dissertation:

An ecological-based approach to examining barriers and facilitators of a physical activity intervention

Committee:

John B. Bartholomew, Supervisor

Keryn E. Pasch

Jessica Duncan Cance

Esbelle M. Jowers

Andrew E. Springer

An ecological-based approach to examining barriers and facilitators of a physical activity intervention

by

Vanessa Leigh Errisuriz, BA; MA

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

August 2016

Dedication

This dissertation is dedicated to my parents and my husband. Your support and encouragement have helped me reach greater heights than I ever thought possible. I love you.

Acknowledgements

I would like to acknowledge my dissertation committee members for their guidance throughout this project. I would also like to acknowledge the members of the Exercise and Sport Psychology Lab who offered their time and support during this process.

An ecological-based approach to examining barriers and facilitators of a physical activity intervention

Vanessa Leigh Errisuriz, PhD

The University of Texas at Austin, 2016

Supervisor: John B. Bartholomew

Texas I-CAN! promotes physical activity (PA) among elementary school children by incorporating 10-15 minute, physically-active, academic lessons into the classroom. A socioecological approach to evaluate effectiveness could provide a deeper understanding of mechanisms promoting or hindering PA. Three studies examined the impact of implementation quality on child PA during active lessons. Teachers from 20 schools self-reported attitude and perceived behavioral control (PBC) related to implementation, and perceptions of school climate. Staff observed teacher feedback to students during active lessons and student PA. Student PA was also measured objectively (i.e. accelerometry). Before examining how teacher-level factors interact to impact student PA, foundational work was necessary. First, several mathematical cut-points have been developed to classify PA intensity among children. Though research indicates that cut-point selection impacts classification of PA among children aged 6-10 years, this has not demonstrated with school-specific PA. Study 1 demonstrated that cut-point selection impacts estimates of in-school PA intensity and students meeting PA guidelines. Second, quality of process (i.e. teachers' ability to engage students in intervention programs) has been linked to program implementation. Study 2 examined associations between teacher feedback during lessons and staff-rated, class PA intensity. Positive associations between PA-related feedback (i.e. reinforcement, technical instruction) and PA intensity were found. Technical instruction was positively associated with how often and how many students were active during lessons. Negative feedback was inversely related to these outcomes. Study 3, then, examined the interrelatedness of quality of process (i.e. PA-related feedback), teacher-level data (i.e. attitudes, PBC, perceptions of school climate), and implementation dose, and their impact on objectively-measured student PA using structural equation modeling. PA-related feedback and dose were positively associated with PA intensity. PBC and attitude towards implementation were positively related to dose. Perception of higher quality school climate was associated with greater PBC and poorer attitudes. PBC was positively, and attitudes negatively, associated with PA-related feedback. Results may inform optimization of future physically-active academic lesson interventions. Identification of factors that impact implementation of active lessons provides opportunities to tailor teacher trainings to focus on these important factors and to intervene if implementation begins to wane during intervention periods.

Table of Contents

List of Tables

List of Figures

Chapter 1: Executive Summary

1.1 Summary Background and Significance

Prevalence of childhood overweight and obesity has dramatically increased over the past 30 years. Data from the 2011-2012 National Health and Nutrition Examination Survey (NHANES) indicate that among children aged 6-11 years, 34.2% are overweight and 17.7% are obese (Ogden, Carroll, Kit, & Flegal, 2014). Childhood obesity is considered a key predictor of adult obesity (Whitaker, Wright, Pepe, Seidel, & Dietz, 1997) and is linked to a variety of adverse physical and mental health problems (Calle & Kaaks, 2004; Daniels et al., 2005; Freedman, Dietz, Srinivasan, & Berenson, 1999; Mauriello et al., 2006; Wilson, 2007). It is not surprising, then, that childhood obesity has become a serious public health concern.

Behavioral factors, including lack of adequate physical activity (PA), contribute to a child's weight status. Federal PA guidelines recommend at least 60 minutes of moderate or vigorous PA per day for children, most of which should be aerobic in nature (2008 Physical Activity Guidelines for Americans). Unfortunately, the percentage of children meeting these recommendations declines as they age (Fakhouri, Hughes, Brody, Kit, & Ogden, 2013). Moreover, physical inactivity is associated with larger gains in body mass index (BMI) (Moore et al., 2003) and increased risk of obesity (Andersen, Crespo, Bartlett, Cheskin, & Pratt, 1998; Dennison, Erb, & Jenkins, 2002; Trost, Kerr, Ward, & Pate, 2001).

Numerous factors are associated with children's PA behavior and weight status, including neighborhood safety (Molnar, Gortmaker, Bull, & Buka, 2004; Singh, Siahpush, & Kogan, 2010), access to recreational facilities (Dunton, Kaplan, Wolch, Jerrett, & Reynolds, 2009; Sallis & Glanz, 2006), and the school PA environment (Davison & Birch, 2001). The school PA environment may be of particular importance, as it can address the aforementioned environmental issues associated with PA. For instance, schools are a relatively safe environment and can provide access to recreational facilities that support PA behavior. Additionally, children spend more time in schools than in any other setting when away from home (Story, Nanney, & Schwartz, 2009). Moreover, schools are an obvious setting where a large number of children, regardless of demographics (e.g. gender, race/ethnicity, socioeconomic status), can be targeted. Therefore, school-based obesity intervention programs have the potential to be highly effective.

Opportunities for PA throughout the school day are limited due to increasing focus on standardized testing. Time spent in recess and physical education class has often been reduced in order to ameliorate pressure associated with achieving high academic standards (Davison & Birch, 2001). As a result, there has been a recent push to explore interventions within the school day that encourage PA outside of the physical education class. Physically-active, academic lessons, which combine PA with teaching of academic content, are a particularly attractive intervention strategy. They incorporate more PA for students throughout the school day without detracting from time spent in academic instruction. Research has demonstrated that physically-active, academic lessons are associated with increases in the amount of daily PA, attentional focus, and decreases in BMI and behavioral issues among children. Thus, the attributes and effects of physically-active, academic lessons may serve to create more buy-in from schools as an intervention.

The finding that physically-active, academic lessons increase children's daily amount of PA is consistent across multiple intervention studies (Bartholomew & Jowers, 2011; Mahar et al., 2006; Stewart, Dennison, Kohl, & Doyle, 2004) and forms of assessment, including pedometers (Bartholomew & Jowers, 2011; Mahar et al., 2006; Stewart et al., 2004), observation (Bartholomew & Jowers, 2011; Donnelly et al., 2009), questionnaires (DuBose et al., 2008) and accelerometry (Bartholomew & Jowers, 2011; Donnelly et al., 2009; Stewart et al., 2004). However, the impact of these lessons on academic outcomes and fitness is less clear. Findings from cross-sectional research support a relationship between PA and academic performance among elementary school children. However, results from randomized and non-randomized prospective studies are mixed (Donnelly & Lambourne, 2011). Discrepancies such as these may be partially due to a failure to take into account quality of implementation, a key variable thought to influence effectiveness of school-based programs (Dusenbury, Brannigan, Falco, & Hansen, 2003). Previous work has found quality of implementation to vary considerably, with programs of higher quality expected to ensure more effective outcomes when disseminated (Dusenbury, Brannigan, Hansen, Walsh, & Falco, 2005). With regard to physically active lessons, Donnelly and colleagues (Donnelly et al., 2009) found that increases in fitness only occurred for those students whose teacher exhibited high implementation rates. If factors that promote or hinder implementation are not considered, then risk of "Type III error" increases, leading one to conclude that the program is not effective when, instead, it was not implemented completely and/or correctly (Dusenbury et al., 2003). It is critical to study how and why a program works, and under which conditions it works, before it can be widely disseminated (Flay et al., 2005).

Though quality of implementation is impacted by several factors, fidelity of implementation, or dosage of program, has been most systematically studied. Other important components of quality of implementation have been largely ignored. For instance, recent research has implicated quality of process (e.g. how well teachers engage students in the program) as an important factor in achieving quality of implementation (Dusenbury et al., 2003, 2005). In addition to quality of process, implementation of school-based, PA programs has been found to be associated with various individual-level factors, including: teacher beliefs and attitudes regarding the compatibility of physically-active lessons to the curriculum (Bartholomew & Jowers, 2011), teacher physical activity (Cothran, Kulinna, & Garn, 2010), perceived barriers to implementation (Bartholomew & Jowers, 2011; Gittelsohn et al., 2003; Singh, Chinapaw, Brug, & van Mechelen, 2009), and perceived behavioral control for implementation (Bartholomew & Jowers, 2011). Surprisingly, there has been minimal consideration of organizational-level characteristics (e.g. school climate) that can affect implementation. However, given the importance of green space, principal support, and other indicators of school climate, it is impossible to understand the effects of individual-level characteristics without considering the environmental context.

Much of the research in this area has neither used nor developed a theoretical model with which to study the relationships between individual-level and school-level factors and their effect on implementation of physically-active academic lessons. Therefore, the focus of this dissertation is to investigate how teacher- and school-level factors influence elementary teacher implementation of physically-active, academic lessons in the classroom, and ultimately, intensity of child PA. The findings of this dissertation should help inform evaluation of similar interventions and contribute to more effective and efficient intervention development.

Before a model can be tested, some limitations of the current research must be addressed. As stated previously, a variety of tools have been used to assess PA among interventions implementing physically-active, academic lessons. Due to limitations associated with self-report measures of PA, accelerometry has become the preferred method by which to objectively assess time spent in various PA intensities among children (Trost, Loprinzi, Moore, & Pfeiffer, 2011). However, in physically active lesson interventions which used accelerometers to assess PA, only a subsample of the participants was assessed. Additionally, it is unclear which cut points were used to determine PA intensity. There are many cut points that have been developed to evaluate intensity of PA among children. Recent work indicates that certain cut points more accurately assess PA intensity among children aged 6 to 11 years old (Trost et al., 2011). Much of this research has been done using PA data which includes time children spend at home, leisure-time, and in-school. Therefore, it is necessary to determine whether the impact of cut point selection remains consistent when only assessing PA during the school day. This will enable the development of a more effective framework for evaluation of school-based program implementation and child PA.

An additional limitation of previous research on this topic is that quality of process, a key component of implementation success, has been understudied and the methodology underdeveloped (Dusenbury et al., 2005). The few attempts to assess quality of process have relied on teacher self-reports (Hallfors & Godette, 2002; Ringwalt et al., 2002) rather than classroom observations (Dusenbury et al., 2005), which can help to limit biases associated with self-report. It is clear that a more objective tool for measurement of this construct must also be developed and assessed before the influence of individual- and school-level factors on implementation can be addressed.

Observation of the type of feedback teachers give students during physically-active, academic lessons may provide a way to more objectively measure how well teachers engage students in the program. Previous work in the physical education setting indicates that the type of feedback teachers provide during class can impact psychological constructs associated with intention to be physically active. For instance, one study found that PE teachers who provided students with positive feedback were more likely to have students who had greater perceived autonomy, competence, and relatedness in regards to PA than those students whose teachers provide more critical feedback (Standage, Duda, & Ntoumanis, 2003). Thus, observation of feedback could be considered a more objective, unobtrusive way to measure quality of process.

Given the significance of developing effective PA interventions, and the influence of implementation quality on intervention success, it is important to identify factors associated with successful implementation. This dissertation aims to investigate the impact of implementation quality on student PA intensity as well as identify teacher- and school-level factors associated with implementation of physically-active, academic lessons by elementary teachers within the classroom. Moreover, this dissertation will address specific limitations of previous work, namely, ensuring the use of appropriate cut points with which to assess PA intensity among elementary school children as well as the creation of a more objective quality of process measure. The findings of this dissertation will help inform development, evaluation, and optimization of future school-based, PA interventions.

1.2 Current Study Aims

1) Compare accelerometer cut points for classifying physical activity intensity during the school day among elementary school children.
2) Examine teacher feedback to students during physically-active, academic lessons as a measure of quality of process (i.e. how well teachers engage students), and determine which types of feedback are associated with greater PA intensity among students.
3) Investigate and test a theoretically-based model with which to study the relationships between teacher- and school-level factors and explain their impact on quality of implementation and child PA.

1.3 Research Purpose

The purpose of this dissertation is three-fold. Study 1 will focus on determining whether the pattern of effects of accelerometer cut point selection on children's PA intensity classification during the school day is consistent with previous research. The objective of Study 1 is to:

1) Examine participants' accelerometer data, obtained during the school day, using various cut points to determine the extent to which classification of PA levels between the various cut points is similar or disparate.

Study 2 will focus on adapting a previously validated observation system, to examine teacher feedback to students during physically-active, academic lessons as a measure of quality of process. Associations between teacher feedback and staff-observed, student PA during active lessons will be explored. Additionally, it is expected that teachers will be grouped into classes of homogeneous profiles as a result of the feedback used during physically-active lessons. The objectives of Study 2 are:

1) Determine underlying teacher feedback profiles through the use of factor profile analysis.
2) Validate teacher feedback profiles through an assessment of their association with student PA.

Study 3 will incorporate the findings of the previous two studies, examining the impact of teacher feedback from Study 2, along with other teacher-level data, including teacher implementation, on student PA – as measured via the cut points from Study 1. The objective of Study 3 is then:

1) To develop and explore a model of the relationships between teacher-level variables and determine their impact on student PA.

1.4 DEFINITIONS OF TERMS

- **Accelerometer:** Objective measurement tools that allow for the estimation of energy expenditure among individuals, as well as the quantification of the amount of time spent sedentary, and in light, moderate, and vigorous physical activity. They are designed to collect and store several days, or weeks, of activity in small sampling intervals (e.g. seconds, minutes) known as epochs. Accelerometers are small, lightweight, and usually worn on the hip, affixed to a belt.
- **Body Mass Index (BMI):** A measure used to determine body fatness based on individuals' height and weight. BMI is calculated as weight in kilograms divided by height in meters squared (kg/m^2), rounded to one decimal place.
- **Child Overweight/Obesity:** For children, overweight is defined as a BMI > 85th percentile, while obese is defined as a BMI > 95th percentile, according to CDC 2000 Growth Charts.
- **Cut Points:** Established through regression equations to describe the relationship between activity counts and EE, which allow cut points to be established (based on counts/minute) for classification of sedentary, light, moderate, and vigorous activity.
- **Type III Error:** A type of error that occurs which leads to concluding a program is not effective when, instead, it was not implemented correctly and/or completely.
- **Metabolic Equivalent or MET:** A unit of measure of the rate at which the body expends energy that is based on the energy expenditure while sitting at rest and is equal to 3.5 milliliters of oxygen per kilogram of body weight per minute

- **Moderate-Intensity Physical Activity:** On an absolute scale, physical activity that is done at 3.0 to 5.9 times the intensity of rest. On a scale relative to an individual's personal capacity, moderate-intensity physical activity is usually a 5 or 6 on a scale of 0 to 10.
- **Physical Activity:** Any bodily movement produced by the contraction of skeletal muscle that increases energy expenditure above a basal level. In these Guidelines, physical activity generally refers to the subset of physical activity that enhances health.
- **Physically Active, Academic Lessons:** Typically, teacher implemented and designed to incorporate 10-15 minutes of moderate-to-vigorous physical activity (MVPA) into the school day, without taking time away from academic instruction. Combine moderate-to-vigorous movement with the teaching or review of academic content.
- **Quality of Process:** How well implementers engage students in an intervention or prevention program
- **Vigorous-Intensity Physical Activity:** On an absolute scale, physical activity that is done at 6.0 or more times the intensity of rest. On a scale relative to an individual's personal capacity, vigorous-intensity physical activity is usually a 7 or 8 on a scale of 0 to 10.

Chapter 2: Literature Review

2.1 Background and Significance

Obesity

In the United States, obesity prevalence rates among all age groups have drastically increased over the past three decades. Consequently, addressing this issue has become the focus of many public health efforts. Obesity can be defined as an excess of body fat (Dehghan, Akhtar-Danesh, & Merchant, 2005), typically resulting from caloric imbalance, and specifically an excess of calories consumed relative to those expended (Daniels et al., 2005). Though there are various ways to measure obesity (e.g. densitometry, magnetic resonance imaging) (Dehghan et al., 2005), most epidemiological studies use Body Mass Index (BMI). BMI is calculated as weight in kilograms divided by height in meters squared (kg/m^2), rounded to one decimal place (Ogden, Carroll, Kit, & Flegal, 2012). For children and adolescents, obesity is defined as having a BMI greater than or equal to the age- and sex-specific 95^{th} percentiles of the 2000 CDC growth charts (Ogden & Flegal, 2010).

Childhood Obesity

Combatting childhood obesity is a significant health priority (U.S. Department of Health and Human Services, 2010) as obesity rates among children are thought to be indicative of future rates among adults. Indeed, obesity in childhood is considered a key predictor for obesity in adulthood (Whitaker et al., 1997). For instance, previous work found children who were obese as early as age 2 were more likely to be obese as adults (Freedman et al., 1999). Childhood obesity has reached epidemic levels, with results from the 2011-2012 National Health and Nutrition Examination Survey (NHANES) indicating that 16.9% of U.S. children and adolescents aged 2-19 years are obese (Ogden et al., 2014). Additionally, prevalence of obesity increases as age increases. Among children aged 2-5 years, 8.4% are classified as obese, compared to 17.7% of children aged 6-11 years, and 20.5% of children aged 12-19 years (Ogden et al., 2014). As the largest increase in obesity prevalence occurs between 2-5 years and 6-11 years, it is important to concentrate obesity prevention efforts among children aged 6-11 years.

Differential Impact of Obesity

Obesity is not equally prevalent across all groups, differentially impacting children of different gender, race/ethnicity, and socioeconomic status (SES). While findings from the most recent NHANES indicate that obesity prevalence is not significantly different between male (16.7%) and female (17.2%) children and adolescents aged 2-19 years, these rates do differ among the various age groups. For instance, among children aged 6-11 years, the discrepancy in obesity rates between males (20.1%) and females (15.7%) is greater than that found between 2-5 year-old males (9.5%) and females (7.2%) and adolescent males (19.6%) and females (17.5%) (Ogden et al., 2014).

Prevalence of obesity also differs significantly by race/ethnicity. The highest prevalence of childhood obesity is among Hispanic children and adolescents (22.4%), followed by their non-Hispanic black (20.2%), non-Hispanic white (14.1%), and non-Hispanic Asian (8.6%) counterparts (Ogden et al., 2014). Finally, findings from epidemiological studies indicate that prevalence of childhood obesity in the United States is higher among low-SES children than those from higher-SES backgrounds (Shrewsbury & Wardle, 2008; Wang, 2001).

Consequences of Childhood Obesity

Childhood obesity is a serious public health concern, linked to a variety of adverse proximal and distal health outcomes, including high cholesterol and hypertension (Freedman et al., 1999), type 2 diabetes (Pinhas-Hamiel & Zeitler, 1996), asthma, orthopedic problems, menstrual irregularities, and sleep apnea (Mauriello et al., 2006; Wilson, 2007), as well as breast and colorectal cancers (Calle & Kaaks, 2004). In addition to physical health, childhood obesity also impacts psychological health. For instance, psychological disorders, such as depression, more frequently occur in obese children (Daniels et al., 2005). If the current obesity epidemic among children is not addressed, it has been speculated that the current generation of children will be the first to have a shorter life expectancy than their parents (Whitaker et al., 1997). As a result, public health efforts remain focused on addressing the issue of childhood obesity.

Determinants of Childhood Obesity

There are many factors associated with the etiology of childhood obesity. While genetics may influence the development of obesity in childhood, a large proportion of the growing epidemic can be contributed to lifestyle factors, such as diet and physical activity behaviors (J. O. Hill & Trowbridge, 1998; Wieting, 2008). Research suggests that obesity in children is associated with greater intake of foods that are energy-dense (Newby, 2007), high in fat (Gillis, Kennedy, Gillis, & Bar-Or, 2002; Guillaume, Lapidus, & Lambert, 1998; Tucker, Seljaas, & Hager, 1997) and carbohydrates (Maillard et al., 2000; Newby, 2007), and greater consumption of sugar-sweetened beverages (Malik, Schulze, & Hu, 2006; O'Connor, Yang, & Nicklas, 2006). However, while obesity typically results from excess energy intake, which implicates poor dietary behaviors, this energy intake must be in excess of energy expenditure.

Federal physical activity guidelines recommend at least 60 minutes of physical activity (PA) per day for children and adolescents, most of which should be aerobic in nature and performed at a moderate or vigorous intensity (2008 Physical Activity Guidelines for Americans). While 76.1% of children aged 6-8 years meet these recommendations, the percentage drops to 64.7% among children aged 9-11 years, indicating the beginning of a steep decline in PA at a young age (Fakhouri et al., 2013). The impact of physical inactivity on risk for childhood obesity is well documented. Cross-sectional studies have demonstrated that children who engage in the least vigorous physical activity are most likely to be overweight or obese (Andersen et al., 1998; Trost et al., 2001). Another study among youth from Mexico City found risk of obesity decreased by 10% for each hour per day of MVPA, while risk increased by 12% for each additional hour per day of television viewing (Dennison et al., 2002). Moreover, a longitudinal study of 103 children found that those who were physically inactive had consistently larger gains in BMI and higher amounts of subcutaneous fat, as measured by triceps and sum of skinfolds (Moore et al., 2003), relative to their more physically active counterparts.

School Physical Activity Environment

Children's environments are thought to play a role in the development of childhood obesity, particularly through their impact on diet and PA behaviors. Of particular interest is the relationship between the school PA environment and children's PA behaviors. Schools provide an obvious setting where children can be exposed to various physical activities (Davison & Birch, 2001). For instance, recess has been associated with increased PA, with previous work finding elementary school children to be engaged in PA 59% of the time during recess, and vigorous activity occurring 21% of the time (Kraft, 1989; Pellegrini & Smith, 1998). Though not directly linked to child obesity, recess is clearly important in obesity prevention as it promotes PA. On the other hand, physical education (PE) classes have demonstrated a direct impact on childhood obesity. Previous work indicates that time in PE has a strong negative effect on BMI change for girls who are overweight or at risk for overweight, where a 1-hour increase in instruction time was associated with a 0.31-point reduction in BMI (Datar & Sturm, 2004). Additionally, longitudinal data have shown that for each weekday that normal weight children and adolescents participated in PE, the odds of becoming an overweight adult decreased by 5 percent (Menschik, Ahmed, Alexander, & Blum, 2008).

However, changes in the school system have negatively impacted the PA environment, providing less opportunity for children to engage in PA and contributing to an obesogenic environment. Over the past 40 years, there has been an increased emphasis on accountability for academic success for both preschool and primary school education (Pellegrini & Bohn, 2005), and consequently, opportunities for PA such as recess and PE have been reduced in many schools in an attempt to ameliorate the pressure related to achieving certain academic standards (Davison & Birch, 2001). Reducing the amount of time children are active in school may also impact PA in general. For example, experimental research has shown children to be less active after school on days when they had no recess and PE class in school (Dale, Corbin, & Dale, 2000), intimating that changing the school PA environment may have far-reaching consequences.

School-Based Physical Activity Interventions

Recently, research efforts have focused on examining the impact of school-based interventions on the development of childhood obesity. One reason for this shift is that schools provide an obvious setting where a large number of children can be targeted to reduce obesity. More than 95% of American youth aged 5 – 17 are enrolled in school; 48 million students attend public elementary, middle, and high schools each day, with an additional 5.3 million students attending private schools (Story et al., 2009). Children spend more time in schools than in any other environment when away from home (Story et al., 2009), approximately 32-33 hours per week (Hofferth & Sandberg, 2001). During the first two decades of life, no other institution has as much continuous and intensive contact, and thus influence on a child (Story et al., 2009). Furthermore, schools can help to address other environmental issues associated with PA. Schools are a relatively safe environment and can provide access to recreational facilities, which supports PA behavior.

Physically Active Academic Lessons

To address the reduction of in-school PA opportunities, innovative approaches, such as the implementation of physically active academic lessons, have been developed in order to increase PA and energy expenditure throughout the day. Physically active academic lessons are typically teacher implemented and designed to incorporate 10-15 minutes of moderate-to-vigorous physical activity (MVPA) into the school day, without taking time away from academic instruction. Several programs have been developed recently that combine moderate-to-vigorous movement with the teaching or review of academic content, including Energizers (Mahar et al., 2006), Take 10! (Kibbe et al., 2011), Physical Activity Across the Curriculum (Donnelly & Lambourne, 2011), and Texas Initiatives for Children's Activity and Nutrition (I-CAN!) program (Bartholomew & Jowers, 2011).

Impact on Physical Activity

Research indicates that physically active academic lessons are effective in increasing the amount, duration, and intensity of PA in elementary school children. The impact of the Take 10! program on children's PA and other related outcomes has been the most frequently studied. Mahar and colleagues (Mahar, Rowe, Kenny, & Fesperman, 2003) found that implementation of Take 10! lessons in 22 classrooms (n=342 students) resulted in a 142 steps/day increase, on average. Another study examined PA levels of 36 children, randomly selected from two elementary schools, during Take 10! activities compared with activity levels during recess, PE, lunch, and afterschool/weekend activities (Moore et al., 2007). PA was monitored with Actical accelerometers over a consecutive 7-day period and the data indicated that PA levels, as indicated by average activity counts per minute, during Take 10! (x=2775.0 counts/min) were significantly higher than those exhibited in PE (x=1813.7 counts/min), lunch (x=1371.5 counts/min), and afterschool/weekend activities (x=1480.4 counts/min) and similar to activity levels during recess (x=2169.7 counts/min) (Kibbe et al., 2011; Moore et al., 2007). The Take 10! program has even demonstrated an impact on children's PA, cross-culturally. An adaption of Take 10! named Happy 10, was piloted in two primary schools (one intervention, one control) in Beijing, China (Liu, Wilson, Qi, & Ying, 2007). Children in the intervention school experienced an increase in PA duration (2.8 hours vs. 3.3 hours) as compared to students in the control school who exhibited a decrease in duration.

Physical Activity Across the Curriculum (PAAC), a modified version of the Take 10! program, has also demonstrated significant and positive impacts on children's PA levels. PAAC combined existing academic lessons taught through PA by classroom teachers in Northeast Kansas with examples from Take 10! activities (Donnelly et al., 2009). PAAC was a three-year, cluster randomized, controlled trial in which 24 elementary schools were randomized to receive PAAC or serve as a control (Donnelly et al., 2009). PA of 4,515 elementary school students (3,465 intervention and 1,050 control) was assessed using the System for Observing Fitness Instruction Time (SOFIT) (Gibson et al., 2008). Higher SOFIT scores indicated higher PA intensity. Results indicated that PA intensity levels of intervention students (3.40) were significantly greater than their control counterparts (2.17) (Gibson et al., 2008; Kibbe et al., 2011). A random subsample of children (n=77 PAAC, n=90 control) were assigned accelerometers and completed four consecutive days of PA monitoring. Findings indicated that children in PAAC schools had significantly greater PA (13%>) and exhibited greater levels of PA during the school day (12%>), on weekends (17%>), and on weekdays (8%>) relative to control students. Children in the PAAC condition also demonstrated significantly greater levels of MVPA (27%>) than children in control schools.

An initial evaluation of the Texas I-CAN! program compared the impact of two school days with and two school days without active lessons on pedometer counts of students in 22 classes from Kinder through 5^{th} grade at a single elementary school. Across all grade levels, I-CAN! lessons led to a significant increase of approximately 1,000 steps, which corresponds to nearly 10 minutes of MVPA (Bartholomew & Jowers, 2011). Another study examined the impact of I-CAN! lessons on child PA across 8 elementary schools (4 intervention, 4 control). Physical activity of students from 47 classes (25 intervention, 22 control) throughout the school day was assessed with pedometers. Students in the intervention condition increased their activity by 300 steps while students in the control condition experienced a reduction in steps by nearly the same amount (Bartholomew & Jowers, 2011). Additionally, a subsample of students wore Actigraph GT1X accelerometers and demonstrated that approximately 20% of the I-CAN! lessons were spent in MVPA (Bartholomew & Jowers, 2011) further indicating that these lessons impact the amount of MVPA students obtain throughout the school day.

Impact on Academic Achievement and Attentional Focus

In addition to their ability to target a large number of children to impact PA and BMI, physically active academic lessons are an attractive intervention strategy because they have been linked to academic achievement and attention control. In the HOPS initiative by (Hollar et al., 2010), a variation of Take 10! was implemented in the classroom and resulted in intervention students achieving significantly higher math scores on the Florida Comprehensive Achievement Test (FCAT) than control students. Moreover, this effect was consistent across all ethnic groups, but particularly apparent among Hispanic children in intervention schools who showed a 20-point increase in FCAT math scores (Hollar et al., 2010; Kibbe et al., 2011). The PASS & CATCH project, which combined Take 10! activities with the Coordinated Approach To Child Health program, examined the association between increased PA during the school day and academic achievement among 932 third and fourth grade students from eight elementary schools (Murray et al., 2008). Students' reading comprehension and math problem solving skills were assessed with Stanford 10 tests at 3 times points over approximately a year-and-a-half. Results indicated that math scores of students in the intervention condition improved significantly more over time relative to controls (Murray et al., 2008). Findings from the three-year PAAC study demonstrated that children in the intervention condition exhibited significant improvements in reading, math, and spelling scores from baseline to three years, compared to control children (Donnelly et al., 2009).

Physically active academic lessons have also demonstrated positive impacts on students' ability to pay attention in class, a key predictor of academic success

(Greenwood, Horton, & Utley, 2002). For instance, Mahar and colleagues (Mahar et al., 2006) examined the impact of Energizers lessons on change in children's on-task behavior before and after the lesson compared to change in on-task behavior before and after traditional academic lessons. Observations of on-task behavior before and after intervention and control lessons were each 30 minutes in duration in which six students were observed for five minutes. Findings showed that children who participated in Energizers lessons exhibited a significant improvement in time-on-task from pre- (70.9%) to post-lesson (79.2%). No significant change in on-task behavior from pre- (71.3%) to post-lesson (68.2%) occurred when an Energizers activity was not performed. Moreover, a stronger effect of Energizers on on-task behavior was seen for those students who were least on-task at baseline, resulting in an increase from 46.1% to 66.0% (Mahar et al., 2006).

Time on task was also assessed in conjunction with the Texas I-CAN! lessons among 97 third-grade students (Grieco, Jowers, & Bartholomew, 2009). Students' on-task and off-task behavior was observed through momentary time sampling, using 5-second intervals, for 15 minutes before and after either an active or control lesson. Results indicated that TOT significantly decreased from pre- (83%) to post-lesson (72%) for inactive, control days, while it increased from pre- (86%) to post-lesson (89%) for active lesson days, though this effect was non-significant. These results were even more apparent among children with greater BMI. Overweight and obese children were only able to allot 58% of their time focused on assigned academic material following a sedentary lesson. However, following an active lesson, these children were able to maintain focus on academic material for 93% of the time following an active lesson (Bartholomew & Jowers, 2011; Grieco et al., 2009). The implication that integrating PA into academic lessons can not only increase PA in children but also positively impact academic achievement and attention control has the potential to motivate teachers to use PA as a teaching tool (Mahar, 2011). Therefore, physically active academic lessons may be an effective intervention strategy to improve physical and cognitive health outcomes.

Importance of Quality of Implementation

While there is sufficient evidence of the positive impact of physically active academic lessons on children's PA and academic outcomes, there are many factors that impact their effectiveness. Previous research suggests that one of the most important determinants of a school-based program's success is the fidelity of its implementation (Durlak & DuPre, 2008; Dusenbury et al., 2003). "Fidelity of implementation refers to the degree to which teachers and other program providers implement programs as intended by the program developers" (Dusenbury et al., 2003). While there is general agreement as to what is intended when research refers to fidelity, a single definition, and thus methodology for measuring fidelity, has yet to be determined. Definitions for fidelity include: 1) strict adherence to methods or implementation that conforms to theoretical guidelines of the program; 2) completeness and dosage of each component of the program; 3) quality of delivery of the program; 4) participant engagement; and 5) the degree to which elements distinguishing one program from another are present or absent (program differentiation) (Dusenbury et al., 2003). In order for the results of an intervention to be accurately interpreted, it must be known what aspects of the intervention were delivered (or not delivered) and the quality of their delivery. If the program is not adequately implemented, negative or null findings can occur. Likewise, if implementers deliver a product that is very different from what was intended, positive impact can be achieved that cannot be replicated by other implementers (Durlak & DuPre, 2008).

With regard to physically active academic lesson interventions, though the target is children's PA behavior, implementation is reliant on teachers to achieve intervention success. Yet, little is known about the impact of implementation on physically active academic lesson outcomes. One study by Donnelly and colleagues (2009) found that teacher delivery of PAAC lessons, specifically how many minutes they spent in PA each week, varied considerably over the course of their three-year study. Teacher reports of the number of minutes of PA that students performed each week ranged from 45 minutes to ≥75 minutes, with nine of 14 schools (~64%) averaging ≥75 minutes per week. In turn, results for change in BMI were influenced by the degree of exposure to PAAC. Overall, there was no significant difference for change in BMI for children in PAAC compared to the those in control schools. However, students from schools that had higher delivery of PAAC lessons per week (≥75 minutes) demonstrated significantly smaller increases in BMI and favorable shifts in BMI categories across the three years of the study compared to students from control schools (Donnelly et al., 2009).

While the majority of studies that include a process evaluation component of their physically active lessons examine teacher implementation rates, the impact of these rates on program outcomes is not considered. For example, an evaluation of the effectiveness of the Take 10! program (Mahar et al., 2003) found that teachers reported implementing Take 10! lessons an average of 4.3±1.2 days per week over a 15-week period. Additionally, the average number of steps per class during Take 10 lessons ranged from 505 to 760 steps – which equates to an average of 14% to 19% of total steps taken during the school day (Mahar et al., 2003). Furthermore, Take 10! lessons allowed the intervention group an additional opportunity to be active during the school day, but children in this group only took 142 more steps per day than their control counterparts (Cohen's d=0.14). The authors suggest that the non-significant increase in PA between intervention and control groups may be due to the fact that teachers substituted recess with Take 10!. Yet, it is also possible that the variation in lesson implementation contributed to the small effect size. It is likely that teachers who implemented Take 10! lessons a greater number of times per week had students who exhibited a greater increase in PA. Further, teachers were asked to lead their class in 10 minutes of PA daily using Take 10! materials, however, length of lessons was not reported. Average length of time could have varied among teachers and may have impacted the number of steps students took throughout the school day.

An examination of the four-year (2003-2007) Delta H.O.P.E. (Healthy Options for People through Extension) project's impact on children's PA levels found that implementation patterns of Take 10! activities for teachers varied considerably over the course of the study (Williams et al., 2008; Kibbe et al., 2011). Indeed, the percentage of teachers integrating three or more Take 10! activities into their classroom per week was 51% the first year. This increased to 68% and 63% the second and third years, respectively, but fell to 45% in the last year of the project (Williams et al., 2008). Interestingly, even though the number of teachers implementing at least three Take 10! lessons per week in their classroom fell below 50% in the final year of the project, teachers that year documented an average of 26.8 minutes of PA per week (Williams et al., 2008; Kibbe et al., 2011). The target outcome of Delta H.O.P.E. was to increase student PA in the classroom by 30 minutes per week, and it is clear that the reported average comes close to this goal, regardless of the rates of implementation. While this discrepancy may be due to a reliance on teacher self-report, the study also fails to take into account the length of each lesson. It is unclear if or how teacher implementation rates were associated with average PA levels.

A study of the original efforts to develop Texas I-CAN! (Bartholomew & Jowers, 2011) showed that few teachers within three elementary schools implemented lessons on a regular basis (<25%). While children within two of the three schools exhibited significantly more steps on intervention days than on control days (Cohen's d=.40, p<.05), it is uncertain whether the amount of steps significantly differed dependent on how often teachers implemented lessons as intended. It is possible that an even greater effect size for step count would be demonstrated for children with teachers who implemented lessons on a daily basis.

It is promising that physically active academic lesson interventions still demonstrate positive impacts on children's PA levels and time on task, even if they are not fully implemented as intended by all teachers in intervention conditions. However, the magnitude of the effect of these types of lessons on child outcomes varies. It is important to investigate possible contributors to intervention success. If factors that promote or hinder fidelity of implementation are not considered, then risk of "Type III error" increases - leading one to conclude that the program is not effective when, instead, it was simply not implemented correctly and/or completely (Dusenbury et al., 2003). Conversely, a program may be considered effective when not implemented as intended (Durlak & DuPre, 2008) – an effect that is unlikely to be replicated. Therefore, it is critical to study how and why a program works, and under which conditions it works before it can be widely disseminated (Flay et al., 2005).

Factors Associated with Quality of Implementation

Previous research has identified factors that may be associated with quality of implementation of school-based PA interventions, including teacher beliefs and attitudes regarding the compatibility of physically-active lessons to the curriculum (Lloyd, Cook,

& Kohl, 2005), teacher physical activity (Cothran et al., 2010), and perceived barriers to implementation (Gittelsohn et al., 2003; Singh et al., 2009). Most literature regarding health programs has focused upon individual-level characteristics of the implementers with minimal consideration of organizational characteristics that can affect implementation (Gittelsohn et al., 2003).

There has been little work examining teacher or school characteristics that may impact fidelity of implementation of physically active academic lessons. For instance, a review of the past ten years of Take 10! (Kibbe et al., 2011) found that though teachers are willing to implement Take 10! activities in their classrooms, the characteristics of those who implement the program to a greater extent are not yet clear. The authors conclude that more research needs to be done to explore factors related to implementation. However, findings from a process evaluation of initial Texas I-CAN! efforts may provide a starting point for future research. This evaluation revealed that while teachers endorsed the concept of physically active lessons, few teachers implemented lessons on a daily basis (Bartholomew & Jowers, 2011). A lack of planning time and available resources (i.e. model lesson plans and equipment) were identified as significant barriers to implementation.

A later incarnation of Texas I-CAN!, implemented in 25 classrooms, determined that implementation fidelity (defined as implementation rates) was associated with teacher ratings of lesson quality ($r = 0.52$), self-efficacy to implement the lessons ($r = 0.47$), and perceived barriers to implementation, such as lack of time ($r = -0.58$ (Bartholomew & Jowers, 2011). The authors suggest that the Theory of Planned Behavior (TPB, Ajzen, 1985), particularly the attitude and perceived behavioral control constructs, may play an important role in teacher implementation rates and suggest that training programs might best be centered on TPB. It is clear that more work is needed for this line of research. If factors that impact teacher implementation of physically active academic lessons can be identified, then this provides an opportunity to tailor teacher trainings to focus on these important factors. Additionally, assessment of these factors could provide opportunities to intervene if implementation begins to wane during intervention periods.

2.2 Limitations of Previous Research

Assessment of Physical Activity

Physical activity has been assessed in a variety of ways across physically active academic lesson interventions. Pedometers (Bartholomew & Jowers, 2011; Mahar et al., 2003), the SOFIT observation system (Donnelly et al., 2009; Gibson et al., 2008), and accelerometry (Bartholomew & Jowers, 2011; Moore et al., 2007) have all been used to measure PA exhibited by elementary children in intervention programs. Accelerometry has become the go-to method for measuring PA in free-living children (Trost et al., 2011). This is primarily in response to limitations associated with self-report measures and direct observation of PA as well as the expense and participant inconvenience of more objective measures of PA, such as heart rate monitoring and doubly labeled water (Puyau, Adolph, Vohra, Zakeri, & Butte, 2004; Trost et al., 2011). Accelerometers are objective measurement tools that allow for the estimation of energy expenditure (EE) among individuals, as well as the quantification of the amount of time spent sedentary, and in light [<3 metabolic equivalents (METs)], moderate (3-5.99 METs), and vigorous (≥ 6 METs) PA (Crouter, Clowers, & Bassett, 2006). They are designed to collect and store several days, or weeks, of activity in small sampling intervals (e.g. seconds, minutes) known as epochs. Accelerometers are small, lightweight, and usually worn on the hip, where they are attached to a belt.

The raw accelerometer signal is not typically used to directly quantify PA; instead, the signal is translated into a metric that is anchored to some biological variable, such as EE (Freedson, Pober, & Janz, 2005). Researchers have developed several regression equations to describe the relationship between activity counts and EE, which allow cut points to be established (based on counts/minute) for classification of sedentary, light, moderate, and vigorous activity (Crouter et al., 2006; Trost et al., 2011). The Actigraph accelerometer model has particularly received attention in research and is one of the most widely used in studies involving children (Trost et al., 2011).

Even though accelerometers have become a popular objective measurement tool for PA assessment, few studies of physically active academic lesson interventions have used this method (Bartholomew & Jowers, 2011; Donnelly et al., 2009; Moore et al., 2007; Stewart et al., 2004). Additionally, those studies which used accelerometry to assess PA, have only done so in a random sample of students participating in the intervention, ranging from n=36 (Moore et al., 2007) to n=200 (Bartholomew & Jowers, 2011). It is unknown how representative these random participants are of the larger sample. It may be necessary to assess PA with accelerometry during these types of interventions among a much larger, more representative sample of children to gain a better understanding of how physically active lessons impact PA.

As stated previously, accelerometry uses mathematical cut points to calculate levels of PA intensity. Unfortunately, there are no widely-agreed upon cut points to classify moderate-to-vigorous PA among youth aged 6-10 years. To date, there are at least six cut points that have been developed to classify PA intensity among children (Evenson, Catellier, Gill, Ondrak, & McMurray, 2008; Freedson et al., 2005; Mattocks et al., 2007; Pulsford et al., 2011; Puyau, Adolph, Vohra, & Butte, 2002; Treuth et al., 2004). Table 1 presents a list of cut points developed by each author, and how they were developed. While the cut points developed by Freedson (2005) have been the most frequently used to assess MVPA among children, recent research indicates that they may not be the most accurate. A study by Trost and colleagues (Trost et al., 2011) found that though the Freedson (2005) and Evenson (2008) MVPA cut points demonstrated significantly better classification accuracy than the Mattocks (2007), Treuth (2004), and Puyau (2002) cut points, only Evenson's cut points provided acceptable classification accuracy for all four levels of PA intensity. Moreover, Evenson's cut points performed better than Freedson's among children aged 10 years and younger (i.e. less likely to misclassify light-intensity activity as moderate, leading to overestimation of time spent in MVPA).

Cut Point Name	Sample	Activities	Cut points (Counts per Minute)
Puyau (2002)	N = 26 14 boys, 12 girls M age = 10.7 years Range = 6-16 years	Walking, running, aerobics, soccer, and lifestyle activities (e.g. playing with toys, playing on the computer, jump rope)	Sedentary: < 800 Light: ≥ 800 Moderate: ≥ 3200 Vigorous: ≥ 8200
Freedson (2005)	N = 80 39 boys, 41 girls M age = 11.3 years Range = 6-18 years	Treadmill walking and running (4.4-9.7 km/hr)	Cut points are age-dependent (for a 9-year-old): Sedentary: ≤ 100 Light: > 100 Moderate: ≥ 906 Vigorous: ≥ 1770
Mattocks (2007)	N = 163 73 boys, 90 girls M age = 12.4 years Range =	Lying, sitting, slow walking, brisk walking, jogging, and hopscotch	Sedentary: ≤ 100 Light: > 100 Moderate: ≥ 3581 Vigorous: ≥ 6130
Evenson (2008)	N = 33 12 boys, 21 girls M age = 7.3 years Range = 5-8 years	Sitting, watching TV, coloring books, slow walk, stair climb, basketball, brisk walking, jumping jacks, and running	Sedentary: ≤ 100 Light: > 100 Moderate: ≥ 2296 Vigorous: ≥ 4012
Pulsford (2011)	N = 53 29 boys, 24 girls M age = none given Range = 7 years	Lying down, sitting, slow walking, brisk walking, jogging, hopscotch, basketball	Sedentary: ≤ 100 Light: > 100 Moderate: ≥ 2240 Vigorous: ≥ 3841

Table 1: Accelerometer Cut Points for Children

In regards to physically active academic lesson interventions, it is unclear as to which cut points were used to assess PA intensity among their participants. However, as Freedson cut points have been the most frequently used to assess MVPA among children (Trost et al., 2011), and some of the studies which used accelerometry were implemented prior to development of later cut points (e.g. Donnelly et al., 2009; Moore et al., 2007; Stewart et al., 2004), it is likely that Freedson cut points were used to classify PA intensity. This is concerning in light of the evidence that indicates use of Freedson cut points may lead to overestimation of time spent in MVPA (Kim, Beets, & Welk, 2012; Trost et al., 2011). Previous work comparing the impact of different cut points on children's time spent in various PA intensities has primarily focused on leisure-time PA. Research is needed to determine whether the discrepancies seen in previous work is consistent in other settings (i.e. in-school PA), as cut point selection may impact whether or not a physically active academic lesson intervention is deemed successful.

Quality of Process

As stated previously, of the many aspects of quality of implementation assessed, dosage and adherence to programs has been most systematically studied among school-based prevention and intervention programs, particularly those related to substance use and mental health (Botvin, Baker, Dusenbury, Botvin, & Diaz, 1995; Dusenbury et al., 2005; Rohrbach, Graham, & Hansen, 1993). While assessing and addressing dosage of intervention programs is important, other components of implementation quality have been largely ignored. Research has implicated quality of process (e.g. how well teachers engage students in the program) as an important factor in achieving high quality implementation (Baker, Sudano, Albert, Borawski, & Dor, 2001; Dusenbury et al., 2003, 2005). Within the field of school-based prevention programs, there is general agreement that those programs implemented in schools, outside of highly controlled research studies or efficacy trials, are not implemented with high quality (Domitrovich et al., 2008; Dusenbury et al., 2005). However, quality of process has been understudied, and the methodology underdeveloped (Dusenbury et al., 2005). Initial attempts to assess quality of process relied on teacher self-reports (Hallfors & Godette, 2002; Ringwalt et al., 2002) rather than classroom observations, which can help to limit biases associated with self-report. More recent studies have attempted to more objectively assess quality of process through research staff-rated observations of teacher program implementation.

One study by Dusenbury and colleagues (2005) assessed quality of process of a school-based, substance use prevention program through observations of teacher-implemented lessons. Observers rated 6 items related to how well lessons were delivered and received: 1) teacher-student interactivity, 2) teacher enthusiasm; 3) teachers' communication of goals/objectives; 4) student engagement; 5) student attentiveness; and 6) students expressing their opinions. A quality of process scale was created and was found to be strongly correlated with adherence to implementation ($r=0.66$, $p<0.05$). However, the fact that teachers were asked to re-teach the sessions that were observed during the study, which might have led them to act differently than they would under normal conditions, and the small sample size of teachers ($n=11$), limits the generalizability of these findings.

Another study evaluated the impact of quality of implementation of the KidsMatter Primary program, an Australian mental health promotion, prevention, and early intervention initiative, on student academic outcomes (Dix, Slee, Lawson, & Keeves, 2012). The study lasted for two years and took place across 100 elementary schools. Findings indicated that there was a significant and positive association between quality of implementation and student academic performance. Specifically, the schools that implemented the program with high quality also demonstrated improved learning outcomes for students relative to those in control schools, equivalent to 6 months more of schooling by Year 7. Moreover, this relationship was significant above and beyond the influence of socioeconomic status. It is likely that schools that were committed to high quality implementation of the KidsMatter program positioned themselves to better support students' mental and academic outcomes (Dix et al., 2012).

Giles and colleagues (2008) investigated teacher's quality of implementation of the All Stars program, which is designed to reduce adolescents' participation in risky health behaviors (e.g. substance use). The All Stars program includes 13 classroom sessions that prescribe interactive activities such as games, classroom discussions, and small group activities. Their work examined how quality of process impacted: 1) students' normative beliefs (acceptability and prevalence of problem behaviors among peers); 2) lifestyle incongruence (realization that high risk behavior is incompatible with one's ideals); 3) commitment to avoid high-risk behaviors; 4) student engagement in the program: and 5) past 30-day substance use (i.e. alcohol, tobacco, and marijuana). In this study, the researchers developed an interactivity measure to assess quality of process, specifically student-teacher transactions. Observers coded and rated various teacher-directed behaviors, including whether they praised and encouraged students, accepted/used student ideas during the lesson, asked questions, shared personal self-disclosures about the lesson topic, and classroom management. Findings showed that teachers' use of certain interactive teaching skills were related to more positive student outcomes. Specifically, student-centered methods, such as accepting students' ideas and asking original, repeated, and probing questions, were associated with improvements in students' idealism and normative beliefs regarding risky behaviors and decreases in student marijuana use.

The findings from these studies suggest that quality of process, that is the extent to which teachers engage students in the intervention program, plays an important role in intervention adherence (Dusenbury et al., 2005) and effectiveness (Dix et al., 2012; Giles et al., 2008). Thus, assessment of this construct may be crucial in order to understand why a program is successful or ineffective. The methodology for this construct is newly emerging and highly varied. To date, no study has examined the impact of quality of process on effectiveness of physically active academic lesson interventions. This dissertation is designed to fill this void.

Theoretical Models

Few studies regarding physically active academic lessons have examined factors that impact implementation. What little work that has been done suggests that lack of planning time and available resources adversely impact teacher implementation of physically active lessons while higher teacher ratings of lesson quality and greater self-efficacy to implement lessons enhance implementation fidelity (Bartholomew & Jowers, 2011). Constructs from the Theory of Planned Behavior (TPB, Ajzen, 1985), (e.g. attitude and perceived behavioral control) have been suggested to play an important role in teacher implementation rates (Bartholomew & Jowers, 2011). School-based intervention research has also implicated organizational characteristics, such as school climate, in implementation fidelity (Gittelsohn et al., 2003). Unfortunately, much more research concerning this issue is warranted.

In addition to identifying important teacher and school characteristics related to implementation fidelity and program success, it is also necessary to examine how these factors are interrelated. Use of an ecological perspective, which recognizes that individuals are located within a broader social context (Stokols, 1996; Green & Kreuter, 1999), has offered useful ways of conceptualizing and designing studies intending to enhance public health (Hawe, Shiell, & Riley, 2009). This model may also be useful for evaluation purposes. Using such an approach to examine the effectiveness of a PA intervention could provide a deeper understanding of the mechanisms that promote or hinder PA, as well as provide information that can be extrapolated to future interventions.

The perspective based on dynamic, ecological-systems emphasizes the interrelatedness among a system's parts. Schools are one such system recognized as dynamic, ecological-systems (Trickett & Birman, 1989). Activity settings, time and space bound patterns of behavior, have been identified as crucial elements of an ecological system (Hawe et al., 2009). In schools, an activity setting could be a parent-teacher meeting, science class, or a physically active, academic lesson. In each setting, people, symbols, physical resources, etc. are all significant features of the setting and interact with one another. For a physically active, academic lesson intervention, teachers and the way they interact with the students during lessons, the perceptions teachers have regarding active lessons, and teacher perceptions about school climate may impact the success and quality of intervention outcomes. With an ecological approach, it is the dynamics of the setting that underlies the individual-level behavior (e.g. child PA) rather than the individual-level attributes (e.g. child attitude towards active lessons, child enjoyment). Social systems, such as schools, play important roles in children's development and in turn, greatly impact PA behaviors and may influence the effectiveness of a PA intervention. Characteristics of teachers (e.g. attitudes, beliefs, perceived behavioral control) who implement the intervention may also play a large role in influencing its effectiveness and quality. Therefore, it is imperative to investigate the impact of teacher-level factors, on children's PA during physically-active, academic lessons.

2.3 Current Study

For this dissertation, a more comprehensive assessment of PA is proposed, with specific interest in factors that impact the level of PA during these lessons. Of particular interest is how teacher implementation impacts student PA and how it interacts with other teacher-level factors to have this impact. There is, however, foundational work that must be completed before addressing this question. First, accelerometry uses mathematical cut points to calculate levels of PA. While there is reasonable agreement on adult cut points, this is not the case for children. As a result, Study 1 will explore and compare the impact of cut point selection on estimates of PA intensity during the school day within our sample of students. Likewise, there is no consensus on coding implementation quality and feedback to children during these lessons. The aim of study 2 is to adapt previous observational assessments to examine teacher feedback provided to students during active lessons and the associations between feedback and student PA outcomes. Finally, Study 3 will incorporate the findings of the previous two studies, examining the impact of the coding system from Study 2, along with other teacher-level data, including teacher implementation, on student PA – as measured via the cut points from Study 1.

Chapter 3: Dissertation Study 1

3.1 Background

Accelerometry has become a popular method for assessing physical activity (PA) among children and adolescents (Trost et al., 2011). This is primarily due to the recall and social desirability biases associated with self-report measures (Evenson et al., 2008) as well as the expensive and burdensome nature of other objective measures of PA such as heart rate monitoring and doubly labeled water. Accelerometers quantify the amount of time spent in sedentary, light, moderate, and vigorous activity with the use of dimensionless activity counts in a specified time interval (e.g. seconds, minutes) known as epochs (Evenson et al., 2008). They are small, lightweight, and usually worn on the hip, where they are attached to a belt. The thresholds for PA intensity are developed using a calibration study. Participants in these studies are asked to perform various laboratory (e.g. treadmill) or field (e.g. playing basketball, skipping) activities while wearing an accelerometer. At the same time, energy expenditure is collected from a criterion measure (i.e. indirect calorimetry). Accelerometer counts are then compared with the criterion measure to establish cut points (Evenson et al., 2008).

Among studies in children and adolescents, the ActiGraph accelerometer is one of the most widely used to assess PA. Consequently, the ActiGraph accelerometer models have received a substantial amount of research attention, and at least six sets of youth-specific ActiGraph cut points have been independently developed and published (Evenson et al., 2008; Freedson et al., 2005; Mattocks et al., 2007; Pulsford et al., 2011; Puyau et al., 2002; Treuth et al., 2004). The availability and application of multiple cut points across various studies have resulted in very different estimates of MVPA. This makes it difficult to compare estimates of children's physical activity across studies as well as establish an accurate estimate of population prevalence rates for physical activity among children (Kim et al., 2012). It is not surprising, then, that recent research has emerged to compare the classification accuracy of published sets of these cut points for children and youth (Kim et al., 2012; Trost et al., 2011).

Trost and colleagues (2011) evaluated classification accuracy of five cut points (Freedson/Trost, Evenson, Treuth, Mattocks, and Puyau) in 206 participants, ages 5-15 years. Participants engaged in 12 different types of activities (e.g. lying down, writing, sweeping the floor, basketball, treadmill walk, run) that varied across 4 physical activity intensities while wearing accelerometers. Energy expenditure (EE) was simultaneously measured with indirect calorimetry as a criterion reference standard. Kappa statistics (K) and area under the receiver operating characteristic curve (ROC-AUC) were used to compare the classification accuracy of each cut point to indirect calorimetry. Results indicated that Evenson and Freedson/Trost cut points demonstrated better classification accuracy across all four PA intensity levels (i.e. sedentary, light, moderate, and vigorous) than the cut points developed by Treuth, Mattocks, and Puyau. Only the Evenson cut points exhibited acceptable levels of classification accuracy across all four PA intensities and were considerably more accurate than the Freedson/Trost cut points among children ≤10 years. Specifically, use of the Freedson/Trost cut points among children ≤10 years resulted in the misclassification of light-intensity activities, such as walking and aerobic dance, as MVPA at a rate of 2.3 times that of children older than 10 years. This was not demonstrated when Evenson's cut points were used.

Interestingly, though the Freedson cut points were observed to misclassify light PA as MVPA, the ROC analyses showed almost identical classification precision estimates of MVPA for children ≤10 years between Evenson and Freedson cut points (Trost et al., 2011; Kim et al., 2012). Consequently, Kim, Beets, and Welk (2012) used accelerometer data from the NHANES (2003-2006) to compare estimates of minutes spent in MVPA using Evenson and Freedson cut points, in the subsample of participants 6 to 10 years. Results indicated that use of the Freedson cut points on the NHANES data set estimated 64 to 124 minutes of MVPA per day, while the Evenson cut points indicated a dramatically reduced amount of time in MVPA per day (47-61 minutes). The authors translate these findings into a percent difference in minutes of MVPA ranging from 38%-80% across children ages 6 to 10 years.

Kim and colleagues (2012) suggest that the discrepancy in the classification precision of Freedson cut points between their findings and those by Trost and collesgues (2011) may be due to the challenges in selecting activities that represent the critical point between light and moderate activity. That is, laboratory methods designed to simulate free-living PA are able to accurately capture movements and corresponding accelerometer counts that are below 2296 (Evenson cut point for moderate PA), yet above 1400-1910 (Freedson cut points for moderate PA). This is especially challenging for children aged 10 and younger. Indeed, half of the 12 simulated activities in the study by Trost et al. (2011) had median counts less than 280 per minute while the other half had median counts above 2072 per minute. Given these challenges, Kim and colleagues (2012) label the range of counts per minute difference between the Freedson and Evenson cut points (~1400 to ~2300) as the “zone of possible misclassification” or ZPM. The ZPM represents that range when a sufficient number of counts per minute are accumulated to be classified as MVPA by Freedson’s cut points but classified as light activity by Evenson’s. Since the median counts per minute of the 12 simulated activities performed in the study by Trost and colleagues (2011) did not fall within the ZPM, it is not surprising, then, that their results indicated no differences in the ROC analysis between the Freedson and Evenson cut points. However, findings from Kim and colleagues' (2012) analysis of NHANES data – which utilize truly free-living patterns of activity - suggest that this misclassification can occur at a considerable rate.

The challenge in simulating free-living, child activity is not surprising. In contrast to adult PA that is marked by extended periods of activity (e.g. walking), children's PA is notoriously difficult to assess because it is play-based and characterized by bursts of short activity, which often lasts for a matter of only seconds, intermingled with similar periods of rest (Riddoch, 2004). It is not clear if this pattern of movement typifies children's PA in all settings. Specifically, patterns of children's in-school PA may be very different than the unplanned, play, or lifestyle PA exhibited during leisure-time. Indeed, previous work indicates that children spend significantly more time sedentary in-school versus out-of-school (Gidlow, Cochrane, Davey, & Smith, 2008). Among children attending primary (elementary) school, only a third of total minutes spent in MVPA over a week was accumulated during the school day. This finding is not surprising as opportunities for PA in elementary school are largely limited to physical education class and recess. Additionally, legislation such as No Child Left Behind (NCLB) has increased the pressure for schools to ensure their students are making adequate progress in reading and mathematics, as federal funding is now dependent upon standardized test scores (Sallis et al., 2012). As a result, time spent in PE and recess is often reduced in order to allocate more time to sedentary instruction for reading, language arts, and math (Center on Educational Policy, 2008; Sallis et al., 2012).

Furthermore, the activity that children generally engage in at school may be more likely to fall below the ZPM. Children primarily engage in a light-activity during the school day, exhibited in-class (e.g. walking to get supplies from supply closet) as well as when they rotate from their classroom to special instruction periods (e.g. art, music), to the cafeteria, or to the restroom. However, even during periods designated for PA, such as PE class or recess, children do not seem to engage in sufficient amounts of MVPA. Fairclough and Stratton's (Fairclough & Stratton, 2005) review of PA accumulated during elementary physical education found that students were only active an average of 37% of a 34-minute lesson, or approximately 13 minutes of PA per lesson. Given the relatively light levels of PA, it may be that the Freedson and Everson cut points converge – as was seen with the simulated activities of Trost (2011).

The impact of different cut points on the calculated percent of time spent in MVPA can have serious consequences for research. While an obvious impact would occur on estimates of children meeting PA guidelines (60 minutes of MVPA every day), the degree of this impact is staggering (Loprinzi et al., 2012). Five sets of children's cut points (Evenson, Freedson, Mattocks, Treuth, and Puyau) were applied to accelerometer data taken from the NHANES 2003-2006. Use of the Mattocks cut points resulted in the lowest prevalence estimate (6.2%) while use of the Freedson cut points resulted in the highest (59.3%). Thus, cut point selection resulted in a nearly 10-fold difference in estimates of children meeting guidelines for PA. These findings are consistent with the results of Trost et al. (2011) and Kim et al. (2012). The National Association for Sport and Physical Education (NASPE) recommends that children spend at least 150 minutes per week in physical education class, with at least 50% of their time (75 min.) spent in MVPA (NASPE, 2004). NASPE also recommends that schools incorporate at least 20 minutes of recess per day, in addition to PE classes (100 min.; NASPE, 2006). This suggests a guideline of 175 minutes of MVPA for in-school PA. It is likely that estimates of children meeting these guidelines for in-school PA are also impacted by cut-point selection. However, this has yet to be studied. A physical activity program or intervention may be judged quite differently based on the cut points used to analyze PA data. Given the importance of school-based PA to child health and the use of schools for PA interventions, it is imperative to compare cut points for children's in-school PA.

Therefore, the purpose of this study is to use a sample of elementary school children to extend previous research by 1) examining and comparing the influence of different accelerometer cut points on estimates of PA intensity during the school day, and 2) examining the impact of cut point selection on estimates of children meeting recommended PA guidelines, both daily (60 minutes of MVPA) and at school (175 minutes of MVPA/week). As the cut points developed by Treuth and colleagues (2004) were calibrated with adolescent girls (aged 13-14), these cut points will not be assessed in the current study of elementary school-based PA. However, an additional cut point for children's PA has recently been developed (Pulsford et al., 2011). This study will be the first to examine and compare estimates of PA derived from the Pulsford (2011) cut points to four other commonly-used cut points for children (i.e. Freedson, Evenson, Mattocks, and Puyau). Additionally, previous studies comparing different cut points on PA intensity classification (Trost et al., 2011; Kim et al., 2012; Loprinzi et al., 2012) have used data from NHANES (2003-2006). Unfortunately, these do not exclusively examine child PA during the school day and are not analogous to the present, school-based data. As a result, NHANES data will not be included in the comparison.

3.2 METHODS

Data for the current study were taken from the Initiatives for Children's Activity and Nutrition (I-CAN!) program. This intervention compared the impact of 10-15 minute physically-active academic lessons on student academic achievement, physical activity, and time-on-task, relative to standard academic lessons. All study protocols were approved by the university's Institutional Review Board. District approval for the I-CAN!

project was sought and obtained from four, central Texas school districts. Once district approval was obtained, research staff approached elementary schools within each district for participation. Thirty schools (n=20 intervention and n=10 control) agreed to participate in the I-CAN! program. Fourth grade teachers and students were targeted for recruitment within each school. Teachers in schools were recruited as a team, however, if at least 75% of the fourth grade team (e.g. 3 out of 4 teachers) consented to participate, then they were included in the intervention program. As a result, 149 fourth-grade teachers (n=99 intervention, n=50 control) were recruited for participation in the I-CAN! program. Both parental consent and student assent were required for student participation. Of the 3,128 students eligible for participation in the study, parental consent was obtained from 2,746 students (87.8%). Of those that had parental consent, only 3 students did not assent to participate, leaving a final sample of 2,743 students. As the purpose of this study is to compare the percent of time children spend in various PA intensities during a typical school day, only the PA data from students in control schools are used.

Participants and Design

Participants were 823 fourth-grade students from 10 elementary schools assigned to the control condition of the I-CAN! project. Students and schools were recruited across three academic years (2012-2013, 2013-2014, 2014-2015) from four central Texas school districts. Data were collected over the course of one school week (five days), during school hours. Schools were randomly assigned to have their students' data collected during the Fall (October-December) or Spring (January-May) semester of the academic year, so teachers could have an opportunity to become familiar with implementing active lessons before data collection. Participants were 51% female, 45.3% non-Hispanic white, and - as a proxy for SES - 30% were eligible for free or reduced lunch. See Table 2 for a full description of participant characteristics.

Variable	N*	M(SD)
Age	695	9.5(0.5)
BMI	578	18.5(3.8)
	N*	**%**
Sex	773	
Females	394	51.0
Males	379	49.0
Race/Ethnicity	752	
Non-Hispanic White	341	45.3
Hispanic	221	29.4
Black	57	7.6
Other	134	17.7
Socioeconomic Status	677	
Not Eligible for Free/Reduced Lunch	474	70.0
Eligible for Free/Reduced Lunch	203	30.0
BMI Fitness Zone	582	
Healthy Fitness Zone	375	64.4
Unhealthy Fitness Zone	207	35.6

* Sample size varies due to missing data on each variable

Table 2: Participant Characteristics (n=823*)

Measures

Demographics

Participant demographic information (i.e. sex, age, race/ethnicity, eligibility for free/reduced lunch, BMI) was obtained through school FITNESSGRAM® data.

Physical Activity

During data collection week, students wore an Actigraph GT3X+ monitor in a belt around their waist, positioned on their right hip. Activity counts were collected in 5-

second epochs to best capture the variability in children's activity. Five different children's cut points were used to classify PA intensity (see Table 3): Freedson, Evenson, Mattocks, Pulsford, and Puyau. PA data were downloaded onto a computer and analyzed with ActiLife software, which can separate the accrued counts by time, allowing for specific examination of time engaged in PA during the school day. The percent of time spent sedentary and in light, moderate, and vigorous activity during the school day was calculated with the ActiLife program. Total time spent in MVPA (minutes) were also obtained from the ActiLife program.

Data Analysis

Statistical analyses were performed using SAS 9.7. A series of one-way, repeated measures ANOVAs were run to compare the impact of cut point selection on PA intensity estimates. As each cut point selection is applied to each participants' data, a repeated measures ANOVA was deemed appropriate. Moreover, the nature of the data lends itself to a 3-level model: (1) student PA intensity, (2) nested within class, and (3) nested within school. As nesting violates the assumption of independent observations, which underlies the general linear model, SAS PROC MIXED (SAS Institute, Inc., Cary, North Carolina) was used to account for nesting (Wolfinger & Chang, 1995). Additionally, time spent (in minutes) in MVPA was calculated for each cut point, and from this, the percentage of participants meeting PA guidelines was assessed. Effect sizes were calculated as Cohen's d statistic.

Cut Point Name	Sample	Activities	Cut Points (Counts per Minute)
Puyau (2002)	N = 26 14 boys, 12 girls M age = 10.7 years Range = 6-16 years	Walking, running, aerobics, soccer, and lifestyle activities (e.g. playing with toys, playing on the computer, jump rope)	Sedentary: < 800 Light: ≥ 800 Moderate: ≥ 3200 Vigorous: ≥ 8200
Freedson (2005)	N = 80 39 boys, 41 girls M age = 11.3 years Range = 6-18 years	Treadmill walking and running (4.4-9.7 km/hr)	Cut points are age-dependent (for a 9-year-old): Sedentary: ≤ 100 Light: > 100 Moderate: ≥ 906 Vigorous: ≥ 1770
Mattocks (2007)	N = 163 73 boys, 90 girls M age = 12.4 years Range =	Lying, sitting, slow walking, brisk walking, jogging, and hopscotch	Sedentary: ≤ 100 Light: > 100 Moderate: ≥ 3581 Vigorous: ≥ 6130
Evenson (2008)	N = 33 12 boys, 21 girls M age = 7.3 years Range = 5-8 years	Sitting, watching TV, coloring books, slow walk, stair climb, basketball, brisk walking, jumping jacks, and running	Sedentary: ≤ 100 Light: > 100 Moderate: ≥ 2296 Vigorous: ≥ 4012
Pulsford (2011)	N = 53 29 boys, 24 girls M age = none given Range = 7 years	Lying down, sitting, slow walking, brisk walking, jogging, hopscotch, basketball	Sedentary: ≤ 100 Light: > 100 Moderate: ≥ 2240 Vigorous: ≥ 3841

Table 3: Accelerometer Cut Points for Children

3.3 RESULTS

Percent of time spent in moderate-to-vigorous physical activity during the school day is shown in Table 4. Findings from the mixed-model, repeated measures ANOVAs revealed a significant main effect for cut point selection on time spent in sedentary [$F(4,4056)=1005.72$, $p<0.001$], light [$F(4,4056)=2799.86$, $p<0.001$], moderate [$F(4,4056)=6260.73$, $p<0.001$], and vigorous [$F(4,4056)=1193.20$, $p<0.001$] PA during the school day. Post hoc analyses were run to compare the differences in percent time spent in sedentary, light, moderate, and vigorous PA between cut points, using a Bonferroni adjustment for multiple comparisons (See Table 4 for means and standard deviations for each intensity level; See Figure 1 for a visual representation of the differences between cut points).

Percent Time in Sedentary

Use of Puyau cut points led to a significantly greater estimation of percent time spent in sedentary during the school day than use of Freedson (*Cohen's* $d=2.19$), Evenson ($d=2.51$), Mattocks (d=2.51), and Pulsford ($d=2.51$) cut points ($p<0.001$). Selection of Freedson cut points also led to significantly greater estimation of percent time spent in sedentary relative to Evenson ($d=0.31$), Mattocks ($d=0.31$), and Pulsford ($d=0.31$) cut points ($p<0.001$). There were no significant differences in estimation of sedentary between Evenson, Mattocks, and Pulsford cut points (See Table 4).

Percent Time in Light PA

Freedson cut point selection led to a significantly lower estimate of light activity as compared to Mattocks ($d=3.89$), Puyau ($d=1.33$), Pulsford ($d=3.56$), and Evenson ($d=3.61$) cut points ($p<0.001$). Puyau cut points also estimated significantly less percent of time in light PA relative to Evenson ($d=2.42$), Mattocks ($d=2.82$), and Pulsford

(d=2.38) cut points (p<0.001). Selection of Mattocks cut points resulted in significantly greater estimation of percent time in light PA than use of Evenson (d=0.52) and Pulsford (d=0.55) cut points (p<0.001). Estimation of light PA was significantly greater for Evenson cut point selection compared to Pulsford cut points (p=0.04; d=0.03).

Percent Time in Moderate PA

Selection of Freedson cut points resulted in significantly greater estimation of percent time spent in moderate PA relative to estimates calculated from Evenson (d=3.79), Mattocks (d=4.38), Pulsford (d=3.79), and Puyau (d=3.98) cut points (p<0.001). Use of Mattocks cut points led to significantly lower estimates of percent time spent in moderate PA compared to Evenson (d=1.76), Pulsford (d=1.76), and Puyau (d=1.11) cut points (p<0.001). Both Evenson and Pulsford cutpoints led to significantly greater estimation of percent time spent in moderate PA than Puyau cut points (d=0.54; p<0.001). There was no significant difference between use of Evenson and Pulsford cut points.

Percent Time in Vigorous PA

Use of Pulsford cut points led to significantly greater estimation of percent time spent in vigorous PA compared to Mattocks (d=1.92), Puyau (d=2.73), Evenson (d=0.19), and Freedson (d=0.19) cut points (p<0.001). Puyau cut point selection resulted in significantly less estimation of vigorous PA compared to Mattocks (d=1.05), Evenson (d=2.71), and Freedson (d=2.71) cut points (p<0.001). Estimation of vigorous PA was greater for both Freedson and Evenson cut points relative to Mattocks (d=1.82) cut points (p<0.001). No significant difference was found between Freedson and Evenson cut points.

Cut Point	Mattocks M(SD)	Puyau M(SD)	Pulsford M(SD)	Evenson M(SD)	Freedson M(SD)
Sedentary	68.9(7.2)[a]	83.9(4.4)[b]	68.9(7.2)[a]	68.9(7.2)[a]	71.1(7.0)[c]
Light	28.3(6.9)[a]	12.6(3.8)[b]	24.7(6.1)[c]	24.9(6.1)[d]	8.6(1.9)[e]
Moderate	2.0(0.8)[a]	3.2(1.3)[b]	3.9(1.3)[c]	3.9(1.3)[c]	18.0(5.1)[d]
Vigorous	0.8(0.6)[a]	0.3(0.3)[b]	2.5(1.1)[c]	2.3(1.0)[d]	2.3(1.0)[d]

a-e = Means in a row without a common superscript letter differ ($p<0.05$)

Table 4: Estimates of Percent Time Spent in Sedentary, Light, Moderate, and Vigorous Activity based on Cut Point Selection

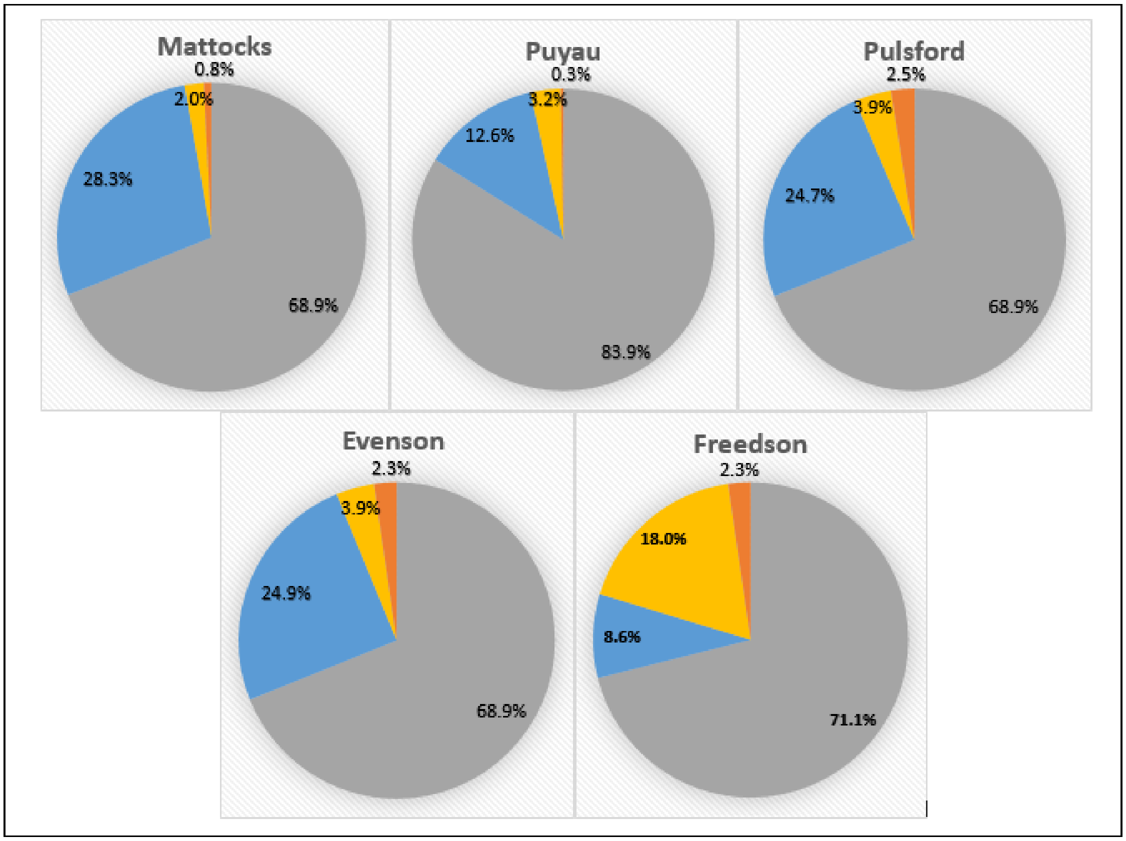


Figure 1: Visual Representation of Differences in Estimated Physical Activity by Cut Point

Meeting Physical Activity Guidelines

Average time spent in MVPA (in minutes) during the school week was calculated for each cut point. Use of Freedson cut points resulted in the highest estimate of time spent in MVPA, followed by Pulsford, Evenson, Puyau, and Mattocks cut points (See Table 5 for means and standard deviations). Percent of students meeting daily PA guidelines (i.e. 60 minutes of MVPA/day or 300 minutes of MVPA/school week) was highest when Freedson cut points were used, whereas no student met daily PA activity guidelines when any other cut point was used (See Table 5). Percent of students meeting school PA guidelines (i.e. 175 minutes of MVPA/school week) was highest when Freedson cut points were used, followed by Pulsford, Evenson, Puyau, then Mattocks (See Table 5).

Cut Point	**MVPA** M(SD)	**% Meeting PA Guidelines**	**% Meeting School PA Guidelines**
Mattocks	52.3(23.6)	0.0	<1.0
Puyau	65.1(27.3)	0.0	<1.0
Evenson	112.6(40.4)	0.0	7.6
Pulsford	116.9(41.6)	0.0	8.8
Freedson	374.3(118.3)	74.0	96.6

Table 5: Weekly Mean MVPA Estimates and Percent of Children Meeting PA Guidelines

3.4 Discussion

Findings from the current study indicate that cut point selection impacts the estimated percent of time that children spend sedentary and in light, moderate, and vigorous PA during a typical school week. For each intensity level, at least one cut point was significantly different from another. In particular, use of Freedson cut points led to significantly greater estimated percent time spent in moderate activity and less percent time in light activity than any other cut point. This is consistent with previous work which found that use of Freedson cut points led to misclassification of light activity as moderate activity (Kim et al., 2012), particularly among children aged 6–10 years (Trost et al., 2011). Additionally, Puyau cut point selection resulted in the greatest percent of time spent in sedentary and the least percent of time spent in vigorous activity, relative to all other cut points. Mattocks cut points resulted in the greatest estimate of percent time spent in light activity and the least percent time spent in moderate activity. This is also consistent with previous work (Loprinzi et al., 2012).

Moreover, this study revealed that the average number of minutes spent in MVPA during the school week was highest when Freedson cut points were used (*M*=374.27) and lowest for Mattocks cut points (M=52.26). Furthermore, use of Freedson cut points resulted in high estimates of prevalence of students meeting daily PA guidelines (74%) and school PA guidelines (96.6%) relative to use of any other cut point. In fact, none of the participants in the study met daily PA guidelines and prevalence of students meeting school PA guidelines ranged from <1% to 8.8% when Evenson, Mattocks, Pulsford, or Puyau cut points were selected. These findings indicate that cut point selection may have a significant impact on prevalence estimates of children meeting PA guidelines, consistent with findings from Loprinzi and colleagues (2012).

This study replicates findings from previous research with a sample of fourth grade students in a specific setting – elementary school. It is worth noting that the estimation of prevalence of students meeting daily PA guidelines resulting from Freedson cut point selection in this study was 74%, compared to the 59.3% found in the study by Loprinzi and colleagues (2012). The difference in prevalence of children meeting daily PA guidelines between the two studies might be attributed to the settings in which the data were collected. In the study by Loprinzi et al., (2012), data were taken from the NHANES (2003-2006) which assesses PA over the course of 7, full days (in and out of school). Children and adolescents were included in their study if they wore the accelerometer at least 4 days, with 10 hours or more of monitoring per day. That is, children's PA during school as well as leisure time PA, outside of school, is measured and analyzed. In the current study, only PA during school was assessed. Since the majority of PA exhibited during school is light intensity, and Freedson cut points tend to misclassify light activity as moderate activity, particularly among children aged 6-10 years, it is not surprising that the current study found a greater prevalence of children meeting daily PA guidelines than previous research.

Unfortunately, this study could not compare estimates of PA intensity derived from cut points to a criterion reference, such as indirect calorimetry. This was not possible in a school setting. Despite the lack of a criterion reference, the results are similar to previous work which did use a criterion reference (Trost et al., 2011; Kim et al., 2012). Based on findings from previous research that compared cut points' PA intensity classification to a criterion reference, it is likely that Evenson's (2008) cut points are the most practical to use when estimating school PA among children aged 6-10 years old. Interestingly, though there were some significant differences found between Evenson and Pulsford cut points, specifically for light and vigorous activity, these differences were marginal – approximately 0.2% for each intensity level, and negligible in magnitude (d=0.03 and d=0.19, respectively). In fact, the difference between these cut points in estimating minutes of MVPA over the school week was about 4 minutes. It is likely that Pulsford's cut points would be a reasonable choice to examine children's PA in school. However, this is the first study, to our knowledge, which examines Pulsford's cut points in comparison with other commonly used cut points. In addition, Pulsford's cut points have yet to be compared to a criterion reference. More research is needed before these cut points should be considered for use in school-based research.

This study is also limited by the sample of participants, derived from elementary schools in central Texas. Though our sample was somewhat heterogeneous (i.e. 51%

female, 45.3% non-Hispanic White, 30% low SES), the PA patterns children exhibit at these schools may not be generalizable. States differ in their physical education and PA policies, and these policies may also differ across districts within the same city. Schools also differ greatly in their environmental support for PA (Harrison & Jones, 2012). Either of these would be likely to impact patterns of activity across different elementary schools. Further research is necessary to describe school PA patterns across a more heterogeneous sample of schools with the goal of demonstrating the impact of these differences in policy and environmental support for PA.

Given the findings of this study and previous research, it is clear that cut point selection exerts a great amount of influence on estimates of PA intensity among children and prevalence of children meeting PA guidelines. Results from school-based, PA intervention studies help to inform public health and public policy decisions. Further research is necessary to determine which cut points provide the most accurate representation of PA patterns in schools.

Chapter 4: Dissertation Study 2

4.1 BACKGROUND

Within the field of school-based prevention programs, there is general agreement that those programs implemented outside of highly controlled research studies or efficacy trials are not implemented with high quality (Domitrovich et al., 2008; Dusenbury et al., 2005). This is concerning as it is well-known that poor implementation is associated with less program effectiveness (Botvin et al., 1995; Durlak, 2015; Payne & Eckert, 2009; Silvia, Thorne, & Tashjian, 1997). As a result, implementation quality, or how well an innovation is carried out in practice (Payne & Eckert, 2009), has become increasingly important. Of the various components of implementation, dosage of (e.g. amount of program content delivered) and fidelity to (e.g. using program content as intended) programs has been most systematically studied, particularly among school-based prevention and intervention programs related to substance use and mental health (Botvin et al., 1995; Durlak, 2015; Dusenbury et al., 2005; Payne & Eckert, 2009; Rohrbach et al., 1993; Silvia et al., 1997). What little research that has examined the impact of implementation on effectiveness of physically active, academic lessons has also focused on dosage and fidelity. While assessing and addressing dosage of physically active, academic lesson interventions is important, other components of implementation quality have been largely ignored.

For instance, research in school-based substance use prevention and mental health promotion programs has implicated quality of process (e.g. how well teachers engage students in the program) as an important factor in achieving high quality implementation (Baker et al., 2001; Dusenbury et al., 2003, 2005). However, quality of process has been understudied, and the methodology underdeveloped (Dusenbury et al., 2005). Attempts to assess quality of process have typically relied on teacher self-reports (Hallfors & Godette, 2002; Low et al., 2014; Ringwalt et al., 2002), which can be biased, rather than classroom observations. As a result, researchers have attempted to more objectively assess quality of process through trained observers of teacher implemented programs.

Most of the observational assessments of quality of process have focused on teacher delivery of the program – particularly how much reinforcement, warmth, or interest teachers provide students, as well as how enthusiastically or clearly they deliver the program content (Goncy, Sutherland, Farrell, Sullivan, & Doyle, 2014). For instance, Dusenbury and colleagues (2005) assessed quality of process of a school-based, substance use prevention by having observers rate 6 items related to how well lessons were delivered and received: 1) teacher-student interactivity, 2) teacher enthusiasm; 3) teachers' communication of goals/objectives; 4) student engagement; 5) student attentiveness; and 6) students expressing their opinions. Results indicated that quality of process was strongly, positively correlated with adherence to implementation.

Giles and colleagues (2008) developed a student-teacher interactivity measure to assess quality of process of the All Stars program, a prevention program designed to reduce risky health behaviors among adolescents. Observers coded and rated various teacher-directed behaviors, including whether they praised and encouraged students, accepted/used student ideas during the lesson, asked questions, shared personal self-disclosures about the lesson topic, and classroom management. Findings showed that teachers' use of student-centered methods, such as accepting students' ideas and asking original, repeated, and probing questions, were associated with improvements in students' idealism and normative beliefs regarding risky behaviors and decreases in student marijuana use.

A similar substance use prevention program implemented with seventh grade children, *keepin' it REAL*, also assessed quality of process as a component of implementation. Observers coded videos of teachers delivering program lessons to their classes (Pettigrew et al., 2014). Quality of process was rated on five dimensions, including teacher attentiveness to students, teacher enthusiasm in delivering content, seriousness in delivering content, the clarity with which the content was provided to students, and positive verbal and nonverbal feedback teachers provided to students throughout the lesson. Consistent with previous work, students who received well-implemented programs showed more positive outcomes (i.e. reductions in substance use) relative to those who received a poorly implemented program.

Of the various components of quality of process, observing the type of feedback teachers provide to students may be particularly relevant in regards to physically active, academic lesson interventions. Previous work in the physical education (PE) setting indicates that the type of feedback teachers provide during class is associated with psychological constructs related to intention to be physically active. One study found that PE teachers who provided students with positive feedback were more likely to have students who had greater perceived autonomy, competence, and relatedness in regards to PA than those students whose teachers provide more critical feedback (Standage et al., 2003). Teacher feedback has also been associated with amount of moderate-to-vigorous physical activity (MVPA) students engage in during PE class time. The System for Observing Fitness Instruction Time (SOFIT; McKenzie et al., 1991) is one commonly used tool to assess teacher behavior during PE class. Of six teacher behavior categories, three are related to type of feedback provided to students during lessons: 1) fitness promotion (i.e. praise/encouragement for PA), 2) class management, and 3) general instruction. Previous work shows positive associations between fitness promotion and percent of time students are engaged in MVPA have been found (Chow, McKenzie, &

Louie, 2008; Martin & Kulinna, 2005) and negative relationships between class management and general instruction and MVPA (Martin & Kulinna, 2005).

One limitation of previous work examining teacher feedback as a quality of process variable is that much of the feedback examined is limited to praise/encouragement, general instruction, and classroom management. However, other types of feedback may be similarly important. In the coaching literature, technical instruction, where the coach instructs the player in a specific technique (i.e. keep your shoulder up), has been linked to positive player outcomes, such as greater self-esteem and enjoyment of sport experience (Smith & Smoll, 1990; Smith & Smoll, 1996), higher perceived ability and competence (Horn, 1985; Weiss, 2000). Thus, it is important to adapt a quality of process measures that incorporates different types of teacher feedback.

Quality of process plays an important role in intervention adherence and effectiveness. Assessment of this construct may be crucial in order to understand why a program is successful or ineffective. However, it is also clear that the methodology for this construct is newly emerging and highly varied. To date, no study has examined the impact of quality of process, specifically teacher forms of feedback, on effectiveness of physically active academic lesson interventions. The purpose of this study was to adapt an observational tool to examine teacher feedback provided to students during physically-active, academic lessons, as a measure of quality of process. Associations between teacher feedback and staff-rated, student PA during active lessons were explored. Relationships between teacher feedback variables would also be examined to determine whether teachers could be organized into groups with homogeneous feedback profiles, which would then be examined in relation to student PA outcomes. It was hypothesized that at least 2 groups, or classes, of teachers would be identified from the data.

4.2 Methods

Data for the current study were taken from the Initiatives for Children's Activity and Nutrition (I-CAN!) program. This intervention compared the impact of 10-15 minute physically-active academic lessons on student academic achievement, physical activity, and time-on-task, relative to standard academic lessons. All study protocols were approved by the university's Institutional Review Board. District approval for the I-CAN! project was sought and obtained from four, central Texas school districts. Once district approval was obtained, research staff approached elementary schools within each district for participation. Thirty schools (n=20 intervention and n=10 control) agreed to participate in the I-CAN! program. Fourth grade teachers and students were targeted for recruitment within each school. Teachers in schools were recruited as a team, however, if at least 75% of the fourth grade team (e.g. 3 out of 4 teachers) consented to participate, then they were included in the intervention program. As a result, 149 fourth-grade teachers (n=99 intervention, n=50 control) were recruited for participation in the I-CAN! program. Both parental consent and student assent were required for student participation. Of the 3,128 students eligible for participation in the study, parental consent was obtained from 2,746 students (87.8%). Of those that had parental consent, only 3 students did not assent to participate, leaving a final sample of 2,743 students.

The purpose of this study was to determine teachers' engagement of their students in I-CAN! lessons from frequency of different types of teacher feedback given to students during I-CAN lessons, as well as to associate these forms of feedback with students' PA levels. Therefore, this study only uses data from teachers in intervention schools who taught I-CAN! lessons (n=94), and staff-rated observations of teachers and students in the intervention conditions (n=100).

Participants and Design

Participants were 94 fourth grade teachers and 1,734 fourth grade students from 20 intervention schools assigned to the intervention condition of the I-CAN! project. Schools, teachers, and students were recruited across three academic years (2012-2013, 2013-2014, 2014-2015) from four central Texas school districts. Teachers were 93.5% female, 84.3% non-Hispanic white, and had a mean age of 38.9 years. Students were 47.4% female, 52.8% non-Hispanic white, and - as a proxy for SES - 37.9% were eligible for free or reduced lunch.

Teacher Training

Once schools and teachers consented to be a part of the study, teachers were invited to attend two trainings sessions: one training in May, at the end of the school year before they were to implement I-CAN! lessons, and a second training in August, prior to the start of the school year in which teachers were to implement I-CAN! lessons in their classrooms. Each training session was between 2 and 3 hours in length. The first session focused on providing teachers with information about the impact of physical activity on academic outcomes. Additionally, teachers were told of the condition of the intervention to which they were assigned (i.e. math lessons, language arts lessons, control) and given information about data collection. During the second sessions, teachers were again briefly provided information and evidence about the link between physical activity and academic outcomes, and procedures regarding data collection were covered more thoroughly. In the second session, intervention teachers were given materials to assist them in implementing academic lessons (i.e. game equipment, ready-made lessons). Teachers assigned to the intervention condition also participated in I-CAN! lessons led by research staff, where the teachers role-played as students.

Data Collection

Parental consent and student assent were obtained for each student. Teacher feedback data and student PA data were collected through staff observations of I-CAN! lessons that were conducted at least once for each class over the course of the school year. Observations took place in late Fall/Winter, between December and February of the school year, to give teachers time to become familiar with implementing active lessons within their classroom.

Measures

Demographics

Teachers reported their sex (i.e. male, female), race/ethnicity (i.e. White, African American or Black, Hispanic, Asian, etc.), and age. Students' demographic information (i.e. sex, age, race/ethnicity, eligibility for free/reduced lunch, BMI) was obtained through school FITNESSGRAM® data and school records.

Teacher Feedback

Teacher feedback was assessed during staff observations of I-CAN! lessons. Categories for teacher feedback were adapted from the Coaching Behavioral Assessment System (CBAS; Smith & Smoll, 1991) and SOFIT (McKenzie et al., 1996): Physical Activity Reinforcement (e.g. "Good high knees!"), Technical Instruction (e.g. "During push-ups, keep your back straight"), Content Instruction (e.g. "72 divided by 9 equals 8"), Content Reinforcement (e.g. "Good answer!"), Game Instruction (e.g. "You should be skipping right now"), Negative Feedback (e.g. "You're doing it wrong"), and Classroom Management (e.g. "Please be quiet"). Observations were completed by two trained research staff, during which they tallied the number of times the teacher gave a particular type of feedback to their students. Counts for each feedback category were summed and averaged across observers.

Research staff that completed observations were trained prior to field observations in various classrooms where I-CAN! lessons were implemented, by the author of the current study, whose ratings were used as the gold standard. During the study, approximately 10% of observations were attended by a third staff member for training purposes. Research staff must have obtained an intra-class correlation (ICC) of at least 0.90 with the gold standard to observe teachers in this study. ICCs are appropriate for assessing inter-rater reliability (IRR) for interval/ratio data and in studies where there are two or more coders (Hallgren, 2012). ICCs for observers ranged from 0.89 for Game Instruction and Classroom Management feedback to 0.98 for Content Reinforcement feedback, indicating excellent inter-rater reliability for teacher feedback variables (Cicchetti, 1994).

Physical Activity

During I-CAN! observations, staff also rated students' PA. Ratings centered on how many children were active [(1) less than half of the class to (3) more than half of the class], how often children were active [(1) not at all to (5) most of the time], the intensity of movement for the lesson [(1) standing still to (5) running]. For each variable, observers' ratings were averaged. The averages were then used in all analyses. ICCs were also calculated to determine IRR among observers for PA variables. ICCs were 0.93 for how often children were active and intensity of movement variables, and the ICC for how many children were active was 0.94, indicating excellent inter-rater reliability for all PA variables.

Data Analysis

Descriptive and Preliminary Analyses

Descriptive and preliminary statistical analyses were conducted with the Statistical Package for the Social Sciences (version 23, IBM, inc.). Frequency distributions of each teacher feedback variable were produced, and means and standard deviations were calculated. Bivariate correlations were run to examine relationships between all feedback variables and student PA.

Latent Profile Analysis

Statistical analyses were performed using MPlus 7.0 (Muthen & Muthen, 2012). A latent profile analysis (LPA) was conducted to examine whether there were subsets or classes of teachers who share similar patterns of responses. That is, LPA was used to determine whether relationships between teacher feedback variables can be used to organize teachers into groups with homogeneous feedback profiles. LPA uses all observations of the continuous dependent variables (i.e. teacher feedback types) to define classes via maximum likelihood estimation (Little & Rubin, 1989). The probability that an individual was properly classified, enabling the individual to be categorized into the best-fitting class, is estimated simultaneously with the overall model (Hill, Degnan, Calkins, & Keane, 2006). Models are estimated with classes added iteratively to determine which model is the best fit to the data. The optimal number of classes for the sample was determined based on several criteria: 1) use of the Lo-Mendell-Rubin Adjusted Likelihood Test (LMRT; Lo, Mendell, & Rubin, 2001); 2) the Akaike information criteria (AIC; Akaike, 1974); and 3) the sample size-adjusted Bayesian information criteria (sBIC; Schwarz, 1978). The LMRT compares the fit of a target model (e.g. 3 class model) to a comparison model that specifies one less class. If the p-value for the LMRT is less than .05, than the solution with more classes is a better fit for the data. The AIC and sBIC are descriptive fit indices where smaller values indicate better model fit. Small classes containing less than 5% of the sample are generally considered spurious, resulting from extracting too many profiles or classes (Hipp & Bauer, 2006). Thus, class size was also considered when determining the optimal number of classes.

4.3 RESULTS

Descriptive analyses indicated that Game Instruction was the most frequently provided feedback during I-CAN! lessons (M = 8.17, SD = 4.62), followed by Classroom Management (M = 6.22, SD = 4.08), Content Reinforcement (M = 4.09, SD = 4.50), Content Instruction (M = 3.59, SD = 3.74), Technical Instruction (M = 2.00, SD = 2.26), Physical Activity Reinforcement (M = 1.65, SD = 2.31), and Negative feedback (M = 1.43, SD = 2.30).

Bivariate correlations revealed that Physical Activity Reinforcement was significantly, positively associated with Technical instruction and students' intensity of movement during I-CAN! lessons. That is, provision of greater amounts of Physical Activity Reinforcement feedback was related to greater frequency of Technical Instruction feedback and higher intensity of movement among students during active lessons. Content Reinforcement was significantly, positively associated with Content Instruction and Classroom Management. Teachers who were observed providing greater amounts of Content Reinforcement feedback to their students during physically-active lessons also provided greater amounts of Content Instruction and Classroom Management feedback. Additionally, positive associations were found between Technical Instruction feedback and how many students were active, how often students were active, and students' intensity of movement during I-CAN! lessons. Conversely, Negative feedback was negatively associated with these three student PA variables. Specifically, greater amounts of Technical Instruction and less Negative feedback provided during lessons was related to a greater number of students active during the lesson, a greater amount of time students spent active during the lesson, and a higher intensity of movement during lessons. No significant associations were found between Game Instruction, Content Instruction, or Classroom Management feedback and any of the student PA variables (See Table 6).

	1	2	3	4	5	6	7	8	9	10
1. Physical Activity Reinforcement										
2. Game Instruction	-.10									
3. Content Reinforcement	.05	-.02								
4. Technical Instruction	.33**	-.14	-.04							
5. Negative Feedback	-.05	.10	-.01	-.08						
6. Content Instruction	.08	-.03	.51**	-.16	.12					
7. Classroom Management	.06	.06	.31**	.01	.10	.18+				
8. How Many Active	.03	-.12	.08	.31**	-.20*	.16	-.01			
9. How Often Active	.05	-.17	-.13	.44**	-.28**	-.10	-.14	.72**		
10. Intensity of Activity	.20*	-.10	-.21	.33**	-.27**	-.01	-.12	.69**	.72**	

Note: * denotes $p \leq .05$; ** denotes $p \leq .01$

Table 6: Bivariate Correlations between all Study Variables (n=100)

Latent profile models containing 2 and 3 classes were fit to the data. The model fit indices for each LPA are available in Table 7. The LMRT indicated that the 3-class solution was not significantly different from the 2-class solution (p = .19). Moreover, the 3-class solution yielded a class size that was too small to be of substantive value (2 teachers for 2% of the sample). Further, the LMRT revealed that the 2-class solution was not statistically different from the 1-class solution, though it approached significance (p = .07). However, the 2-class solution also yielded a class size that was too small to be of substantive value (2 teachers for 2% of the sample). Therefore, a 1-class model was considered the best fit to the data, suggesting that only one homogeneous profile or class represents the data. Teachers in the one class are represented by relatively high levels of Game Instruction and Classroom Management feedback, moderate levels of Content Reinforcement and Content Instruction feedback, and low levels of Negative, Technical Instruction, and Physical Activity Reinforcement feedback.

Solution	LMRT (p)	AIC	sBIC
1 class		3655.92	3648.18
2 class	64.41 (.07)	3605.76	3593.59
3 class	62.69 (.19)	3557.37	3540.78

Note: LMRT = Lo-Mendell-Rubin Test, AIC = Akaike Information Criterion, sBIC = sample size adjusted Bayesian Information Criterion

Table 7: Model fit indices (n=100)

4.4 Discussion

Quality of process is an important, yet understudied and underdeveloped construct, that has been linked to intervention program success. Much of the previous research relies on self-report data from teachers to determine how well teachers engage students in school-based intervention programs. This suggests a need for more objective measures of quality of process. The purpose of this study was to adapt an observation tool to examine teacher feedback to students during physically-active, academic lessons, as a measure of quality of process. Associations between teacher feedback and staff-rated, student PA during active lessons were examined to determine which types of feedback were indicative of a greater number of children active, more time spent being active, and higher intensity of activity. It was thought that relationships between teacher feedback variables would organize teachers into groups with homogeneous feedback profiles, which would then be examined in relation to student PA outcomes.

In the present study, LPA was used to identify groups of teachers with similar feedback profiles. However, the data did not indicate the existence of multiple feedback profiles. Indeed, all teachers were represented by relatively high levels of Game Instruction and Classroom Management feedback, moderate levels of Content Reinforcement and Content Instruction feedback, and low levels of Negative, Technical Instruction, and Physical Activity Reinforcement feedback. As Texas elementary students are given academic instruction and physical activity/physical education instruction from separate instructors (i.e. classroom versus physical education teachers), it is not surprising that classroom teachers provided feedback primarily related to classroom management and academic content during physically active lessons, with little instruction on physical activity per se. Given these findings, it seems reasonable to conclude that most teachers provide similar feedback for physically active, academic lessons in the classroom. This, however, does not indicate that their self-selected instructional style is ideal for achieving high levels of activity.

Though teachers could not be classified into groups based on the type of feedback they provided during active lessons, several feedback categories were individually and directly related to staff-rated observations of student PA during lessons. Specifically, greater Physical Activity Reinforcement and Technical Instruction were associated with higher intensity of movement. Greater Technical Instruction feedback was also associated with how many students were active and how often students were active during I-CAN! lessons. Despite these relationships, these categories were among the lowest frequency of use. That is, most teachers did not utilize Physical Activity Reinforcement or Technical Instruction as a part of their instruction. Conversely, Negative feedback was inversely related to how many students active, how often students were active, and the intensity of activity. Fortunately, few teachers utilized Negative feedback in their instruction. Previous research in the physical education setting indicates that the type of feedback (e.g. positive or critical) teachers provide during class can impact psychological constructs associated with intention to be physically active (Standage et al., 2003), fitness promotion, and percent of time students are engaged in MVPA (Chow, McKenzie, & Louie, 2008; Martin & Kulinna, 2005). In conjunction with previous research, this study demonstrates the importance of assessing forms of teacher feedback, as a measure of quality of process, in achieving program success.

Moreover, these findings provide an opportunity for enhancing teacher training of physically active, academic lessons by emphasizing greater levels of physical activity-related feedback and lower levels of negative feedback. Teacher trainings should incorporate strategies such as modeling and role play to illustrate the importance of providing Physical Activity Reinforcement and Technical Instruction feedback during active lessons. Additionally, it may be beneficial to provide teachers with mentors, that is, teachers who are successful implementers of PA-related feedback during active lessons. These mentors can provide other teachers with different approaches they can use in the classroom to improve quality of process, and specifically the forms of feedback used during active lessons. It is imperative that future research identify which strategies are most impactful with regard to helping teachers improve quality of process of physically active, academic lessons.

In addition, future research should more thoroughly examine the relationship between these teacher feedback categories and student PA during physically active, academic lessons, particularly focusing on whether these types of feedback are associated with more objective measurement of PA (e.g. via accelerometry). It may also be important to identify factors related to delivery of particular types of feedback. For instance, teachers' confidence in their ability to implement the lessons (Perceived Behavioral Control) may contribute to the frequency with which they deliver feedback related to students' PA, such as Physical Activity Reinforcement or Technical Instruction. Additionally, the importance that they place on academic achievement may impact the frequency with which they give feedback related to academic content. Future work should also investigate the factors that contribute to teachers' provision of specific types of feedback.

The null findings of the LPA to identify classes of teachers may be due, in part, to inadequate power to detect the underlying, latent profiles. Although a study by Tein, Coxe, and Cham (2013) found little impact of sample size on power to detect large effects (i.e. Cohen's $d \geq .80$), the smallest sample size used was $n = 250$, more than two times the sample size of the current study. However, there was little power to detect the correct number of profiles for small ($d = .20$) or even medium ($d = .50$) interclass differences. The sample size of the current study was above the limit for the minimum sample size ($n = 70$) for latent class model analyses (such as LPA) as suggested by Wurpts and Geiser (2014). Moreover, Tein and colleagues (2013) estimated 3- and 5-class models, with the smallest cell size of $n = 50$. Thus, it is likely the current study had enough power to assess a 2-class model. An additional limitation to this study is that the number of indicators may not have been sufficient to power the LPA. A greater number of indicators is associated with greater power (Tein et al., 2013; Wurpts & Geiser, 2014), and it is suggested that there be at least 10 indicators of class membership (Tein et al., 2013). It must be noted that statistical power and sample size requirements are under-studied in LPA, and more work is needed in order to understand the sample characteristics and requirements in studies that employ LPA (Tein et al., 2013).

Moreover, though 100 observations of physically active, academic lessons were conducted, feedback was recorded in real-time. Due to the context of the lessons (e.g. outside vs. inside; students in one large group vs. spread out in small groups), it could be difficult to be near teachers at all times and hear all feedback provided during the lesson. It is possible that observers failed to record important teacher feedback. Furthermore, feedback classification was limited to seven categories. Previous research has implicated other types of feedback as important for PA engagement that were not assessed in the present study. Both descriptive (i.e. feedback that describes an error without evaluation) and prescriptive feedback (i.e. feedback that describes how to make performance better) have been linked to increased PA skill performance among middle school students (Silverman, Tyson, & Krampitz, 1992).

Further, the context in which a type of feedback is given (i.e. after a mistake or after a successful performance) may impact PA. Nicaise and colleagues (2007) found that girls perceived greater competence in PA when given higher frequencies of positive feedback after success and less criticism after mistakes. It is likely that assessing a wider range of feedback types as well as when feedback is given could provide greater discrimination between teachers. Future research assessing teacher feedback during physically active, academic lessons should use video recordings to better capture the full range of feedback, and the context in which it is provided, by teachers to their students.

Teacher reactance to observation is another potential limitation of this study. To prevent this effect, teachers were given instruction to implement an active lesson as they normally would and told that observations were focused on students' enjoyment of physically active, academic lessons. Nevertheless, the presence of observers could have impacted teachers' behaviors, and the feedback they provided to students, during lessons. Use of multiple observations over the school year, or multiple video recordings, could help further ameliorate reactance to observation as well as determine whether feedback use is consistent across time.

Finally, this was a cross-sectional design, and so a causal direction between teacher feedback and student PA cannot be determined. That is, whether students are more physically active due to the types of feedback that the teachers are providing or whether teachers are providing greater amounts of physical activity-related feedback because students are moving more cannot be ascertained. Future work in this area should design interventions to examine the causal impact of specific types of feedback on student PA levels.

Results from school-based, PA intervention studies help to inform public health and public policy decisions. It is important to assess the quality of implementation of such interventions, so as to better design teacher training to maximize their impact. The findings of this study suggest that it may be prudent to train teachers to provide specific types of feedback to students during physically active, academic lessons to help promote engagement in PA. However, these data are derived from research staff reports of activity. It is important to assess these categories of feedback on objective measures of student physical activity. This is the basis of Study 3 in this dissertation.

Chapter 5: Dissertation Study 3

5.1 Background

The finding that physically-active, academic lessons increase children's daily amount of PA is consistent across multiple intervention studies (Bartholomew & Jowers, 2011; Mahar et al., 2006; Stewart, Dennison, Kohl, & Doyle, 2004) and forms of assessment, including pedometers (Bartholomew & Jowers, 2011; Mahar et al., 2006; Stewart et al., 2004), observation (Bartholomew & Jowers, 2011; Donnelly et al., 2009), questionnaires (DuBose et al., 2008) and accelerometry (Bartholomew & Jowers, 2011; Donnelly et al., 2009; Stewart et al., 2004). However, previous work among school-based programs has found quality of implementation to vary considerably, with those of higher quality expected to ensure more effective outcomes when disseminated (Dusenbury et al., 2005). With regard to physically active lessons, Donnelly and colleagues (Donnelly et al., 2009) found that increases in fitness only occurred for those students whose teacher exhibited high implementation rates. If factors that promote or hinder implementation are not considered, then risk of "Type III error" increases, leading one to conclude that the program is not effective when, instead, it was not implemented completely and/or correctly (Dusenbury et al., 2003). It is critical to study how and why a program works, and under which conditions it works, before it can be widely disseminated (Flay et al., 2005).

Though quality of implementation is impacted by several factors, fidelity of implementation and dosage of program have been most systematically studied. Other important components of quality of implementation have been largely ignored. For instance, recent research has implicated quality of process (e.g. how well teachers engage students in the program) as an important factor in achieving quality of implementation (Dusenbury et al., 2003, 2005). Unfortunately, this construct is understudied within research focusing on physically active, academic lessons. For example, research among general school-based, PA programs has found implementation to be associated with various individual-level factors, including: teacher physical activity (Cothran et al., 2010) and perceived barriers to implementation Gittelsohn et al., 2003; Singh, Chinapaw, Brug, & van Mechelen, 2009), with no examination of quality of implementation. Likewise, a review of ten years of Take 10! (Kibbe et al., 2011) found that though teachers are willing to implement Take 10! activities in their classrooms, the characteristics of those who implement the program to a greater extent and with a higher quality are not yet clear. The authors conclude that more research needs to be done to explore factors related to implementation.

Findings from a process evaluation of initial Texas I-CAN! efforts may provide a starting point for future research. While teachers endorsed the concept of physically active lessons, few teachers implemented lessons on a daily basis (Bartholomew & Jowers, 2011). A lack of planning time and available resources (i.e. model lesson plans and equipment) were identified as significant barriers to implementation. A later incarnation of Texas I-CAN!, implemented in 25 classrooms, determined that implementation fidelity (defined as implementation rates) was associated with teacher ratings of lesson quality ($r = 0.52$), self-efficacy to implement the lessons ($r = 0.47$), and perceived barriers to implementation, such as lack of time ($r = -0.58$ (Bartholomew & Jowers, 2011). The authors suggest that the Theory of Planned Behavior (TPB, Ajzen, 1985), particularly the attitude and perceived behavioral control constructs, may play an important role in teacher implementation rates and suggest that training programs might best be centered on TPB.

Surprisingly, there has been minimal consideration of teachers' perceptions of organizational-level characteristics (e.g. school climate, principal support, quality of space, etc.) that can affect implementation. Social systems, such as schools, may greatly influence the effectiveness of a PA intervention not only through a direct impact on children's behaviors, but also through their impact on teacher delivery of the program. Given the potential importance of indicators of school climate, it is impossible to understand the effects of teacher characteristics without considering the environmental context in which they occur. Furthermore, much of the research in this area has neither used nor developed a theoretical model with which to study the relationships between teacher characteristics and their effect on implementation of physically active academic lessons. Use of an ecological perspective, which recognizes that individuals are located within a broader social context (Stokols, 1996; Green & Kreuter, 1999), has offered useful ways of conceptualizing and designing studies intending to enhance public health (Hawe et al., 2009). This approach may also be useful for evaluation purposes.

The perspective based on dynamic, ecological-systems emphasizes the interrelatedness among a system's parts. Schools are one such system recognized as dynamic, ecological-systems (Trickett & Birman, 1989). Activity settings, time and space bound patterns of behavior, have been identified as crucial elements of an ecological system (Hawe et al., 2009). In each setting, people, symbols, physical resources, etc. are all significant features of the setting and interact with one another. Active lessons could be considered activity settings. For a physically active, academic lesson intervention, teachers and the way they interact with the students during lessons, the perceptions teachers have regarding active lessons, and teacher perceptions about school climate may impact the success and quality of intervention outcomes. With an ecological approach, it is the dynamics of the setting that underlie the individual-level behavior (e.g. child PA) rather than the individual-level attributes (e.g. child attitude towards active lessons, child enjoyment). Using such an approach to examine the effectiveness of a PA intervention could provide a deeper understanding of the mechanisms that promote or hinder PA, as well as provide information that can be extrapolated to future interventions.

One variable that is frequently overlooked in implementation science research is quality of process – how well teachers engage students in the intervention program. In study 2, the feedback that teachers provided during physically active, academic lessons was observed and investigated as a component of quality of process. Specifically, this study examined whether specific types of feedback (e.g. feedback related to physical activity, content, classroom management, etc.) were associated with greater intensity of class physical activity. Results indicated that higher frequency of Physical Activity Reinforcement and Technical Instruction feedback was associated with greater intensity of class PA. However, study 2 utilized a subjective measure of PA via staff observations. It is important to determine whether PA-related feedback is associated with more objective measures of PA (i.e. accelerometry). Certain teacher characteristics, such as PBC and attitudes towards implementing lesson may also be associated with PA-related feedback, impacting the frequency with which certain types of feedback are provided.

More work is needed for this line of research. If factors that impact teacher implementation of physically active academic lessons can be identified, then this provides an opportunity to tailor teacher trainings to focus on these important factors. Additionally, assessment of these factors and their interrelationships could provide opportunities to intervene if implementation begins to wane during intervention periods. It is imperative to investigate the impact of teacher characteristics on class PA. The focus of this study was to investigate how teacher-level factors influence implementation of physically-active, academic lessons in the classroom, and ultimately, intensity of PA performed by their classes.

5.2 METHODS

Data for the current study were taken from the Initiatives for Children's Activity and Nutrition (I-CAN!) program. This intervention compared the impact of 10-15 minute physically-active academic lessons on student academic achievement, physical activity, and time-on-task, relative to standard academic lessons. All study protocols were approved by the university's Institutional Review Board. District approval for the I-CAN! project was sought and obtained from four, central Texas school districts. Once district approval was obtained, research staff approached elementary schools within each district for participation. Thirty schools (n=20 intervention and n=10 control) agreed to participate in the I-CAN! program. Fourth grade teachers and students were targeted for recruitment within each school. Teachers in schools were recruited as a team, however, if at least 75% of the fourth grade team (e.g. 3 out of 4 teachers) consented to participate, then they were included in the intervention program. As a result, 149 fourth-grade teachers (n=99 intervention, n=50 control) were recruited for participation in the I-CAN! program. Both parental consent and student assent were required for student participation. Of the 3,128 students eligible for participation in the study, parental consent was obtained from 2,746 students (87.8%). Of those that had parental consent, only 3 students did not assent to participate, leaving a final sample of 2,743 students.

The purpose of this study is to identify important teacher characteristics related to implementation quality of physically-active academic lessons, as well as examine how these factors are interrelated and impact program success (increasing child MVPA). This study will incorporate the findings of the previous two studies, examining the impact of PA-related feedback variables from the coding system from Study 2, along with other teacher-level data, including teacher implementation, on student PA – as measured via the cut points from Study 1. Consequently, this study only used data from teachers and students in the intervention condition.

Participants and Design

Participants were 94 fourth grade teachers and 1,734 fourth grade students from 20 intervention schools assigned to the intervention condition of the I-CAN! project. Students and schools were recruited across three academic years (2012-2013, 2013-2014, 2014-2015) from four central Texas school districts. Teacher consent was obtained and parental consent and student assent were obtained for each student. Teacher characteristics (i.e. attitude towards implementation, perceived behavioral control, and perception of school climate) were assessed prior to lesson implementation in August and at the end of the school year in May, respectively. Teacher feedback data and student PA data were collected through staff observations conducted at least once a year. Observations took place in late Fall/Winter, between December and February of the school year, to give teachers time to become familiar with implementing active lessons within their classroom. Students' accelerometer data were collected over the course of one school week (five days). Schools were randomly assigned to have their students' data collected during the Fall (October-December) or Spring (January-May) semester of the academic year. Teachers were 93.5% female, 84.3% non-Hispanic white, and had a mean age of 38.9 years. Students were 47.4% female, 52.8% non-Hispanic white, and – as a proxy for SES – 37.9% were eligible for free or reduced lunch.

Measures

Demographics

Teachers reported their sex (i.e. male, female), race/ethnicity (i.e. White, African American or Black, Hispanic, Asian, etc.), number of years teaching at present school, number of years teaching fourth grade, and number of years teaching overall. Students' demographic information (i.e. sex, age, race/ethnicity, eligibility for free/reduced lunch, BMI) was obtained through school FITNESSGRAM data.

Teacher Characteristics

Teacher *attitude toward implementation* was assessed with 6 items (Cronbach's $\alpha = 0.91$) that ask teachers to rate how they feel about implementing the I-CAN! program for 15 minutes a day, four days a week. Items are bipolar adjective pairs (e.g. unpleasant – pleasant; useless – useful; worthless – valuable) and were scored on a 7-point, Likert scale (from 1 to 7) and summed. *Perceived behavioral control* was assessed with 12 items (Cronbach's $\alpha = 0.92$) that ask teachers whether they are confident that they can implement the Texas I-CAN! program in the face of a series of constraining factors (e.g. preparing for standardized tests, inclement weather, etc.) along with their confidence that a series of facilitating factors will be present (e.g. easily incorporated into curriculum, easily modified to fit with curriculum, etc.). Items were scored on a 7-point Likert scale from 1 (Strongly Disagree) to 7 (Strongly Agree) and summed.

Teacher's *perception of school climate* was assessed via 23 items from three subscales adapted from the Organizational Climate Description Questionnaire for Elementary Schools (OCDQ-RE; Hoy, Tarter and Kottkamp, 1991). These subscales assess school climate by examining teachers' perceptions of principal behavior along three dimensions: 1) the extent to which it is *supportive* (e.g. principal shows a genuine concern and respect for teachers; Cronbach's $\alpha = 0.95$), *directive* (e.g. principal demonstrates rigid monitoring and control of teaching behavior; Cronbach's $\alpha = 0.89$), and *restrictive* (e.g. the principal burdens teachers with paperwork or committee requirements that interfere with teaching; Cronbach's $\alpha = 0.80$). All items were rated on a

4-point Likert-type scale from 1 (rarely occurs) to 4 (very frequently occurs). Items for directive and restrictive subscales were reverse coded to ensure that higher scores for each scale indicate higher quality school climate (i.e. more supportive, less directive, and less restrictive). While these responses are often averaged across participants to provide an overall school climate score, the goal was to assess individual teacher behavior. As a result, a summed score for each subscale was calculated for each teacher.

PA-Related Feedback

Teacher feedback was assessed during staff observations of I-CAN! lessons. During observations, trained staff tallied the number of times the teacher gave a particular type of feedback to their students: Physical Activity Reinforcement (e.g. “Good high knees!”), Technical Instruction (e.g. “During push-ups, keep your back straight”), Content Instruction (e.g. “72 divided by 9 equals 8”), Content Reinforcement (e.g. “Good answer!”), Game Instruction (e.g. “You should be skipping right now”), Negative Feedback (e.g. “You’re doing it wrong”), and Classroom Management (e.g. “Please be quiet”). Frequency of feedback was summed and averaged across all observers for each feedback category. Categories for teacher feedback were adapted from the Coaching Behavioral Assessment System (CBAS; Smith & Smoll, 1991) and SOFIT (McKenzie et al., 1991).

Research staff that completed observations were trained prior to field observations in various classrooms where I-CAN! lessons were implemented, by the author of the current study, whose ratings were used as the gold standard. During the study, approximately 10% of observations were attended by a third staff member for training purposes. Research staff must have obtained an intra-class correlation (ICC) of at least 0.90 with the gold standard to observe teachers in this study. ICCs are appropriate for assessing inter-rater reliability (IRR) for interval/ratio data and in studies where there are two or more coders (Hallgren, 2012). ICCs for observers ranged from 0.89 for Game Instruction and Classroom Management feedback to 0.98 for Content Reinforcement feedback, indicating excellent inter-rater reliability for teacher feedback variables (Cicchetti, 1994).

For the purpose of this analysis, PA-related feedback, (i.e. Physical Activity Reinforcement and Technical Instruction) were used since they were found to be significantly, positively correlated with staff-rated observations of student PA during I-CAN lessons ($r = 0.20$, $p < 0.05$ and $r = 0.33$, $p < 0.01$, respectively) in study 2.

Dose of Implementation

Teachers self-reported the number of times per week they implemented an I-CAN! lesson on a custom made checklist (see Appendix A) where they wrote down the title of the lesson for that day (i.e. Math Freeze Tag). The total number of lessons implemented across the school year was calculated for each teacher.

Physical Activity

Students' PA data was collected via accelerometer on a separate week. During data collection week, students wore an Actigraph GT3X+ monitor in a belt around their waist, positioned on their right hip. Activity counts were collected in 5-second epochs to best capture the variability in children's activity. PA data were downloaded onto a computer and analyzed with ActiLife software, which can separate the accrued counts by time, allowing for specific examination of time engaged in PA during the school day. The percent of time spent in moderate-to-vigorous physical activity (MVPA) during a physically active, academic lesson was calculated using Evenson (2008) cut points. For each class, an average of the percent of time students spent in MVPA during the lesson was calculated for purpose of analysis.

Data Analysis

Statistical Analyses were performed using SPSS (version 23, IBM, inc.) and Mplus 7.0 (Muthen & Muthen, 2012). A structural equation model was used to examine the interrelatedness of teacher and school characteristics as well as assess the impact that each had on student PA outcomes through teacher implementation (See Figure 2). SEM was selected because it provides a number of advantages over path analyses. The model depicted in Figure 2 shows hypothesized relationships between variables, some of which are non-recursive in nature. The issue of non-recursive relationships between variables violate assumptions of path analysis (Keith, 2014), and consequently SEM is the preferred technique. In addition, they hypothesized model is based on a number of latent variables, e.g. school climate. A latent variable is unobservable or unmeasured (Keith, 2014) and calculated from a number of related indicators. In calculating latent variables, SEM also estimates measurement error in a way that is not provided with path analysis. Given multiple indicators, SEM can therefore account for measurement error in the model, thereby getting a more accurate indication of the construct of interest. SEM does this via a confirmatory factor analysis of the measured variables (to calculate the latent variables of interest). Confirmatory factor analysis evaluates the adequacy of individual items as indicators for the hypothesized latent constructs (Lei & Wu, 2007). The "fit" of the model to the data will be considered good if the following criteria are met: 1) the Root Mean Square Error of Approximation (RMSEA) is $\leq .05$, 2) the Standardized Root Mean Square Residual (SMSR) is $\leq .05$, and 3) the Comparative Fit Index (CFI) is $\geq .95$, as suggested by Hu and Bentler (1999).

In traditional SEM, all variables and indicators are assumed to be independent across units; however, in multilevel settings this assumption is violated as units are nested in clusters (i.e. teachers within schools) (Rabe-Hesketh, Skrondal, & Zheng, 2007). While a multilevel SEM would be appropriate for use with these data, the sample size for clusters (in this case, schools) is required to be at least 40 if the between-level model is simple and more than 100 to detect smaller effects at the between level (Meuleman & Billiet, 2009). Unfortunately, the sample size for clusters in these data is n = 20 schools, below the minimum requirement. Additionally, all variables entered into the model are teacher-level characteristics. Due to these limitations, use of a multilevel SEM would be underpowered and could lead to inaccurate results. In response, this study used a single-level structural equation model that controlled for school assignment for each teacher.

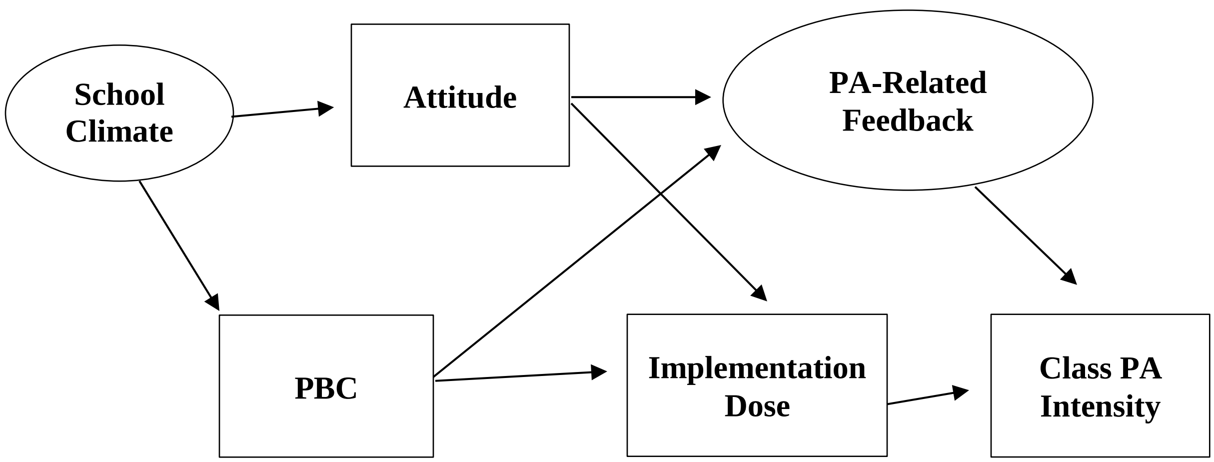


Figure 2: Theoretical Model for Study 3

5.3 RESULTS

Descriptive statistics were run using SPSS software. Means, standard deviations, and correlations between all study variables included in the model are presented in Table 8. Sample sizes for each variable ranged from n = 75 to n = 99 due to missing data. Missing data resulted from participants failing to respond to items on questionnaires or, in the case of feedback variables, staff being unable to observe teachers during active lessons. Of particular concern were the missing data for the PA-Related Feedback and School Climate latent variables in the model (i.e. Physical Activity Reinforcement and Technical Instruction feedback, as well as the three principal behavior subscales, respectively). Indeed, initial attempts to fit the hypothesized model to the data resulted in nonconvergence due to missing data on these indicators and the small sample size. Several datasets were created in order to obtain the largest sample size with which to fit the hypothesized model: 1) Participants were included if they were not missing all three indicators for School Climate (n = 93); 2) Participants were included if they were not missing both indicators for PA-Related Feedback (n = 86); 3) Participants were included if they were not missing all indicators for School Climate and PA-Related Feedback (n=100); and 4) Participants were included only if they had a complete set of data (n = 71). Chi-square analyses (Table 9) and independent-samples t tests (Table 10) were run to examine whether demographics of participants missing data differed from those with complete sets of data within each dataset. Significant differences in age and number of years teaching at their current school were found between participants missing data for school climate indicators and those not missing data ($p < 0.05$). These differences were not significant for those missing data for PA-Related Feedback indicators. No other differences were found. As there was only one participant missing data for both School Climate and PA-Related Feedback indicators, missing data analysis was not conducted for this data set.

	1	2	3	4	5	6	7	8	9
1. Lesson MVPA									
2. PA Reinforcement	-0.04								
3. Technical Instruction	0.19^{+}	0.39^{**}							
4. Number of Lessons/Year	0.28^{**}	-0.04	0.27^{*}						
5. Attitude	0.06	0.13	0.20^{+}	0.21^{*}					
6. PBC	-0.04	0.12	0.09	0.30^{**}	0.54^{**}				
7. Directive Behavior	0.22^{*}	0.13	0.20^{+}	0.17	0.15	-0.03			
8. Restrictive Behavior	0.09	0.11	-0.01	0.05	0.17^{+}	0.12	-0.12		
9. Unsupportive Behavior	-0.02	-0.09	0.11	0.12	-0.26^{*}	-0.32^{**}	0.10	0.04	
Mean (SD)	22.57 (8.31)	1.64 (2.33)	2.05 (2.28)	126.11 (26.67)	32.39 (6.27)	64.86 (12.34)	15.18 (3.86)	28.80 (6.05)	16.20 (6.05)

Note: * denotes $p \leq .05$; ** denotes $p \leq .01$

Table 8: Bivariate correlations, means, and standard deviations of all study variables (n = 75 to n=99)

Missing Data on School Climate Indicators						
	Gender (%)		Race (%)		Ethnicity (%)	
	Male	Female	White	Non-White	Hispanic	Non-Hispanic
Cases with complete Data (n = 93)	6.5%	93.5%	92.4%	7.6%	11.2%	88.8%
Cases missing data (n = 7)	0.0%	100.0%	100.0%	0.0%	28.6%	71.4%
p-value	0.49		0.45		0.18	

Missing Data on PA-Related Feedback Indicators						
	Gender (%)		Race (%)		Ethnicity (%)	
	Male	Female	White	Non-White	Hispanic	Non-Hispanic
Cases with complete Data (n = 86)	5.8%	94.2%	94.1%	5.9%	10.8%	89.2%
Cases missing data (n = 14)	7.1%	92.9%	85.7%	14.3%	23.1%	76.9%
p-value	0.85		0.28		0.22	

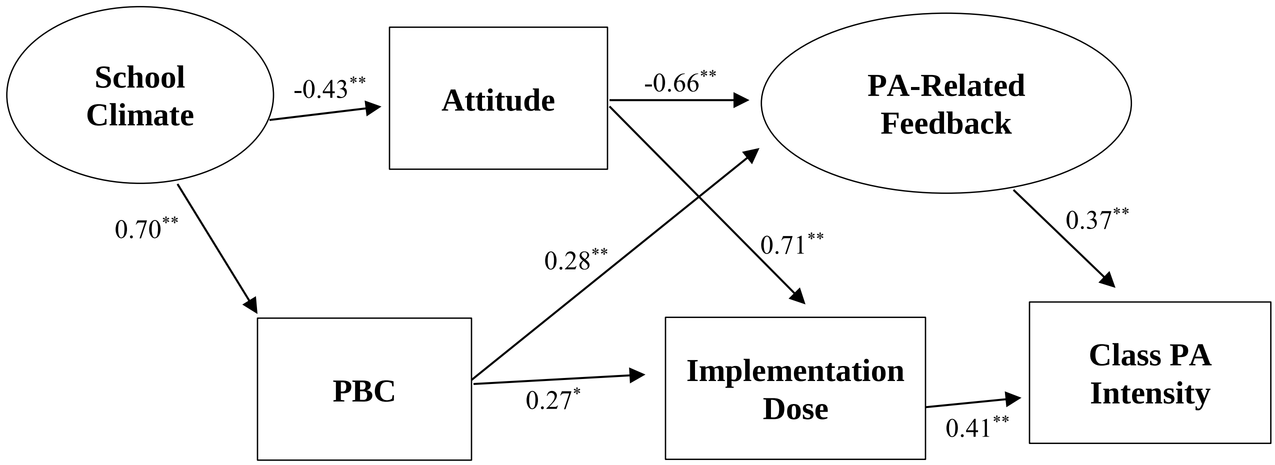
Table 9: Comparison of characteristics between participants missing data and those with complete data with a χ^2 test

Missing Data on School Climate Indicators								
	Age		#Years Teach 4th Grade		#Years Teach Overall		#Years Teach at Current School	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Cases with complete Data (n = 93)	38.17	9.33	4.53	4.31	9.39	6.19	4.02	3.57
Cases missing data (n = 7)	46.00	9.18	3.57	4.89	9.86	6.72	6.86	2.55
p-value	**0.04**		0.58		0.85		**0.04**	

Missing Data on PA-Related Feedback Indicators								
	Age		#Years Teach 4th Grade		#Years Teach Overall		#Years Teach at Current School	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Cases with complete Data (n = 86)	38.93	9.49	4.53	4.47	9.69	6.42	4.37	3.74
Cases missing data (n = 14)	37.43	9.28	4.00	3.49	7.79	4.39	3.29	2.23
p-value	0.59		0.67		0.29		0.29	

Table 10: Comparison of characteristics between participants missing data and those with complete data with an independent samples *t* test

The hypothesized model was able to converge when only teachers with a complete set of data were included in the model (n=71). Results for the hypothesized model for class PA intensity were RMSEA = 0.06, SMSR = 0.06, and a CFI = 0.96, indicating relatively good model fit. The standardized coefficients for the significant paths at the 0.05 level are shown in Figure 3. Results indicated that perception of school climate was negatively related to attitude towards implementing lessons. That is, teachers who perceived a higher quality school climate – more supportive, less directive and restrictive – reported poorer attitudes towards implementing active lessons. Conversely, school climate was positively associated with perceived behavioral control regarding lesson implementation, indicating that teachers who perceived higher quality school climates reported greater PBC in implementing active lessons in the face of facilitating and constraining factors. Attitude towards implementing lessons was negatively associated with PA-related feedback, with more positive attitudes towards implementation indicative of lower frequencies of PA-related feedback. PBC was positively associated with PA-related feedback, with more confidence in ability to implement lessons associated with greater amounts of PA-related feedback provided during active lessons. Both attitudes and PBC were positively related to implementation dose. Specifically, higher attitude and PBC scores were associated with greater number of lessons implemented over the school year. Finally, number of lessons implemented per year and PA-related feedback were significantly and positively associated with class PA intensity during physically active, academic lessons. Students whose teachers provided them with greater numbers of lessons over the school year and greater amounts of PA-related feedback during active lessons showed higher percentage of time spent in MVPA during active lessons.



Note: *p < 0.05, **p < 0.01

Figure 3: The Relationship between Teacher Characteristics, Quality of Process, Implementation Dose, and Class PA Intensity

5.4 DISCUSSION

This study examined the interrelatedness of specific teacher characteristics and investigated their influence on implementation of physically-active, academic lessons, and ultimately, class PA intensity. A hypothesized model was tested using structural equation modeling. Teacher characteristics were associated with one another as well as with implementation of physically active, academic lessons and class PA intensity. The primary outcome to this intervention is focused on the percent of time the class spent in MVPA. Both PA-related feedback and the number of lessons implemented over the school year were significantly and positively associated with class PA intensity. The more PA-related feedback teachers provide during physically active lessons and the greater number of lessons implemented over the school year, the more time spent in MVPA during an active lesson. The association between number of lessons implemented over the school year and percent of time in MVPA during active lessons is consistent with previous work. Dose has consistently been associated with high quality implementation of school-based programs. Thus, teachers who are more compliant meeting the daily recommendation of active lessons are also more likely to be compliant with the goal to include a higher percent of time in MVPA during each lesson. The relationship found between PA-related feedback and percent of time spent in MVPA has the potential to be of greater impact. If teachers engage in more PA-related feedback it appears to result in greater MVPA. This provides a point of emphasis for training that relates to the primary outcome of higher intensity physical activity.

This study also demonstrated that both PBC and attitude towards implementation were associated with the number of active lessons that teachers implemented throughout the year. Teachers with more positive attitudes about the lessons and greater confidence that they could implement lessons in the face of barriers and facilitators, implemented a greater number of lessons over the course of the school year. This is consistent with previous research with physically active, academic lessons that found implementation rates to be related to teacher beliefs and attitudes regarding compatibility of lessons to the curriculum and PBC (Bartholomew & Jowers, 2011). Findings from the present study are also consistent with implementation research among school-based, substance use prevention programs that link fidelity of implementation to more positive teacher attitudes about the program and confidence in their ability to teach the interactive methods (Beets et al., 2008; Parcel et al., 1995).

Components of a school's climate are thought to influence those who come into contact with it (Gittelsohn et al., 2003). Based on the diffusion of innovations model, it is posited that administrators influence teacher characteristics such as attitudes and PBC through a social process where a program is evaluated (favorably or unfavorably) through the perceptions of the social system in which they are embedded (Beets et al., 2008; Gittelsohn et al., 2003; Rogers, 2002). In this study, perception of school climate was significantly associated with PBC. Specifically, perception of a higher quality school climate – less directive, restrictive, and more supportive – was associated with greater PBC among participants. In turn, teachers who reported greater levels of PBC exhibited higher PA-related feedback (i.e. teachers gave more physical activity-related feedback to students during lessons). This is also congruent with previous research in the physical education setting that demonstrates that a lack of confidence in ability to teach PE is related to poor quality programs and adversely related to children's capacity to achieve PE outcomes (DeCorby, Halas, Dixon, Wintrup, & Janzen, 2005; Morgan & Hansen, 2008; Xiang, Lowy, & McBride, 2002). Regular classroom teachers, who focus on academic content, may not feel confident in their ability to provide PA-related feedback during active lessons, and thus, not provide as much of this type of feedback. On the other hand, those teachers that report higher PBC may complete higher quality lessons. The positive relationships between school climate and PBC, and PBC and quality of process, are consistent with previous work regarding implementation. These findings indicate the importance of targeting PBC during teacher trainings.

The relationships between school climate and attitudes, and PA-related feedback and attitudes, in the current study are neither intuitive nor consistent with previous research. Beets and colleagues (2008) found that a supportive school climate was indicative of more positive attitudes towards implementation of a school-based, substance use prevention program. In the present study, perception of a higher quality school climate – less directive, restrictive, and more supportive – was associated with poorer attitudes toward implementation of lessons. However, various elements of school climate may be differentially associated with attitude towards implementation. Beets and colleagues (2008) only examined perceptions of school climate in terms of support. The current study used a latent construct of school climate that included three components: directive, restrictive, and supportive. Findings may reflect that teachers have positive attitudes towards implementing something novel in their classroom. Restrictive school climates have been linked to lower perceptions of autonomy among teachers. Teachers are less likely to see that they have choices in selecting actions that lead toward positive outcomes if principals do not engage in empowering behaviors (Davis & Wilson, 2000). Moreover, less perceived autonomy among teachers has been associated with job stress and dissatisfaction (Davis & Wilson, 2000; Pearson & Moomaw, 2005). Teachers who perceive a poorer school climate may have more favorable attitudes towards any innovation, including physically active lessons, as it allows them more autonomy within the structure of their curriculum. Further research is necessary to elucidate the impact of specific components of school climate on attitudes regarding implementation of physically active lessons.

More positive attitude towards implementation was related to lower PA-related feedback. This inverse relationship is somewhat surprising. It was thought that this association may be due to teachers' beliefs that physically active lessons improve academic outcomes. If a teacher's positive attitudes regarding active lessons stems from the belief that it will help students achieve academically, then it is likely that they would limit PA-related feedback while increasing amount of feedback related to academic content. A bivariate correlation between teacher attitudes and feedback did not show significant associations between attitude and content encouragement ($r = .06$, $p > .05$) or content instruction feedback ($r = .04$, $p > .05$). This suggests that attitudes, while important in terms of the number of lessons implemented over the year, are not indicative of the amount of PA-related feedback provided during such lessons. Previous work in the PE setting suggests that while teacher beliefs and attitudes about the importance of PE was an asset, a number of factors served as barriers to implementation that negatively impacted the quality of outcomes (DeCorby et al., 2005). Therefore, emphasizing confidence in ability to implement physically active lessons during teacher trainings rather than attitudes may be more effective in impacting frequency of PA-related feedback. More research is needed to elucidate the relationship between attitudes towards physically active lessons and feedback given during lessons.

The primary limitation to this work is the small sample size. This may lead to error in parameter estimations, bias, and reduced power. However, previous research has advanced various rules-of-thumb for sample size, requiring 5 or 10 observations per estimated parameter (Bentler & Chou, 1987; Bollen, 1989), which this study meets. Additionally, more recent work has demonstrated that models with stronger effects generally required fewer cases, although this effect is nonlinear (Wolf, Harrington, Clark, & Miller, 2013). The small sample size also contributed to an inability to examine models with a greater number of relationships posited between variables to see if they were better fitted to the data due to issues with convergence. There was no assessment of quality of process beyond the amount of PA-related instruction. Fortunately, the hypothesized model fit the data relatively well. As the findings of the present study are consistent with similar work examining implementation of school-based programs, there is strong confidence in the variables assessed, though they must be interpreted with caution until other variables can be tested directly.

Another limitation, related to sample size, was the inability to examine a multilevel model. The nested nature of the data (i.e. teachers within schools), lends itself to a hierarchical model. Unfortunately, the high rate of missing data on predictor variables, which led to approximately one-third of participants being excluded from the model, in addition to the inadequate number of clusters available (n=20) did not allow for multilevel modeling. It must be noted that school was controlled for in the analysis, and there were few differences between the model when controlling for school relative to when school was not controlled.

Teacher reactance to observation may be another study limitation. To prevent this effect, teachers were given instruction to implement an active lesson as they normally would and told that observations were focused on students' enjoyment of physically active, academic lessons. Nevertheless, the presence of observers could have impacted teachers' behaviors, and the feedback they provided to students, during lessons. Use of multiple observations over the school year, or multiple video recordings, could help further ameliorate reactance to observation as well as determine whether feedback use is consistent across time.

The cross-sectional nature of the data also limits ability to interpret findings and establish causal effects. Future studies could be designed to train different components of quality of process to determine causal influence. In addition, accelerometer data was not collected at the same time as the observational data. As a result, it is not possible to provide a temporal link with PA-related feedback. However, examination of the hypothesized model with staff-rated PA intensity was very similar to the model described in the current study, with only one path – from number of lessons implemented per year to PA intensity – not reaching statistical significance, though it was a similar effect (See Appendix B for full model). Moreover, use of accelerometer data from a second week ensures an outcome that is independent of the validation outcome from study 2 (i.e. staff-rated class PA intensity). For instance, any spurious relationships between feedback categories and PA outcomes from study 2 are minimized with the use of accelerometer data. Additionally, any reactance to research staff presence during observation of active lessons is avoided through the use of accelerometer data – collected during lessons when no researcher was present. Therefore, relationships found within the model, though cross-sectional, can be interpreted with some confidence.

The purpose of this study was to enhance knowledge of factors that impact quality of implementation of physically active, academic lessons. Despite the aforementioned limitations, the present study provides some insight into teacher characteristics that impact quality of implementation of physically active, academic lessons, and in turn, child PA intensity. Consistent with previous work, PA-related feedback and dose of implementation were both associated with more positive program outcomes - in this instance, classrooms with higher PA intensity. Greater amounts of PA-related feedback provided to students during active lessons was related to higher PA intensity during lessons. Teacher trainings for these lessons should, in part, focus on instructing teachers about the best types of feedback to give during lessons. Future research should investigate the various training strategies (e.g. model a lesson, role play, etc.) to increase use of this type of feedback.

Attitudes towards implementing active lessons and PBC both positively contributed to implementation dose – the number of lessons implemented over the course of the year. However, only PBC was positively associated with PA-related feedback. This suggests that while attitudes towards a program are important in creating buy-in for and implementing a program, confidence in ability to implement lessons may be of greater importance as this was associated with higher quality lessons. Additionally, PBC was adversely impacted by school climate. This is not surprising given the additional burdens placed on teachers from more restrictive school climates. To improve teachers' PBC, it may be beneficial to obtain greater administrative support for implementation of both the present program and other activities in school. This general support may reduce competing burdens that will, in turn, increase PBC for active lessons. This might also provide teachers with resources and/or skills necessary to overcome barriers to high quality implementation and instill greater confidence in their ability to implement physically active, as well as all other, lessons.

Chapter 6: Discussion

Due to the high prevalence of childhood overweight and obesity and the adverse physical and psychological consequences associated with this health concern, there is a need for public health efforts to focus on the development and implementation of effective obesity prevention and intervention programs. The percentage of children meeting physical activity (PA) recommendations declines as they age (Fakhouri et al., 2013) and physical inactivity is associated with larger gains in BMI (Moore et al., 2003) and increased risk of obesity (Dennison et al., 2002; Trost et al., 2001). It is unsurprising, then, that school-based obesity prevention programs have become a popular means of intervention as a large number of children, regardless of gender, race/ethnicity, or SES, can be targeted. Implementing physically-active, academic lessons within the classroom has become a particularly attractive intervention strategy. These lessons are able to increase amount of daily PA without detracting from time spent in academic instruction. This is especially important considering that other opportunities for PA during the school day, such as physical education and recess, are limited due to increased pressure associated with achieving high academic standards (Davison & Birch, 2001).

The implementation quality of physically-active, academic interventions and the factors that contribute to their quality are understudied. It is critical to understand how and why a program works and under which conditions it works before it can be disseminated (Flay et al., 2005). Much of the research in this area has neither used nor developed a theoretical framework with which to study the relationships between teacher characteristics and their effect on implementation of physically-active academic lessons. Therefore, the focus of this dissertation was to investigate how teacher factors influence elementary teacher implementation of physically-active, academic lessons in the classroom, and ultimately, intensity of PA exhibited by their class. Before this could be examined, other limitations of current research in this field needed to be addressed.

For instance, in most physically-active, academic lessons intervention research, PA is assessed with pedometers (Bartholomew & Jowers, 2011; Mahar et al., 2003) and observation (Donnelly et al., 2008; Gibson et al., 2008), while only subsets of students were assessed with accelerometry (Bartholomew & Jowers, 2011; Moore et al., 2007) – the preferred method by which to objectively measure time spent in various PA intensities (Trost et al., 2011). More importantly, there are a variety of cut points that have been developed to evaluate PA intensity among children aged 6 to 11 years old, and it is unclear which cut points in previous work were used to assess PA intensity. Further, recent research has indicated that some cut points have better classification accuracy. Evenson (2008) cut points in particular have been demonstrated to have acceptable classification accuracy across all four levels of PA intensity among children aged 6 to 10 years (Trost et al., 2011). This was in contrast to use of Freedson (2005) cut points – the most widely used cut point for children – that overestimated children's engagement in MVPA. Previous work comparing the impact of different cut points on children's time spent in various PA intensities has primarily focused on leisure-time PA. However, results from school-based, PA intervention studies help to inform public health and public policy decisions. Therefore, study 1 of this dissertation was designed to determine whether the discrepancies seen in previous work is consistent in other settings (i.e. in-school PA), as cut point selection may impact whether or not a physically active academic lesson intervention is deemed successful.

Results from study 1 demonstrated comparable findings to previous research comparing the impact of cut point selection on estimates of child PA. Freedson (2005) cut points led to the greatest estimated percent time spent in moderate activity and less percent time in light activity than any other cut point. Moreover, this study revealed that the average number of minutes spent in MVPA during the school week (*M*=374.27) and estimates of prevalence of students meeting daily PA guidelines (74%) and school PA guidelines (97.7%) was highest when Freedson cut points were used. Study 1 extends the literature by replicating findings from previous work with a sample of children within a specific setting – elementary school.

Though this study was limited in that estimates of PA intensity derived from cut points could not be compared to a criterion reference, such as indirect calorimetry. Despite this, the results are similar to previous work that did use a criterion reference (Trost et al., 2011; Kim et al., 2012). Moreover, based on findings from previous research that compared cut points' PA intensity classification to a criterion reference, it is likely that Evenson's (2008) cut points are the most practical to use when estimating school PA among children aged 6-10 years old. Future assessments of child PA should focus on validating the use of Evenson (2008) cut points within the school setting to determine which provide the most accurate representation of in-school PA patterns. Given the findings of this study and previous research, it is clear that cut point selection exerts a great amount of influence on estimates of PA intensity among children and prevalence of children meeting PA guidelines.

Additionally, before a theoretical model for program implementation could be assessed, it was necessary to develop a more objective assessment of quality of process (e.g. how well teachers engage students in the program), a key component of implementation success (Dusenbury et al., 2005). The few attempts to assess quality of process have relied on teacher self-reports (Hallfors & Godette, 2002; Ringwalt et al., 2002) rather than classroom observations (Dusenbury et al., 2005), which can help limit biases associated with self-report. Previous research in the physical education setting indicated that the type of feedback (e.g. positive or critical) teachers provide during class can impact psychological constructs associated with intention to be physically active (Standage et al., 2003). Observation of feedback could be considered a more objective, unobtrusive way to measure a component of quality of process of physically-active, academic lessons. Thus, study 2 was designed to adapt an observational coding system to derive teacher feedback to students during physically-active, academic lessons, as a measure of quality of process.

Study 2 examined associations between trained observer ratings of teacher feedback and student PA during active lessons. Results indicated that several feedback categories were individually and directly related to staff-rated observations of student PA during physically active, academic lessons. Greater amounts of Physical Activity Reinforcement and Technical Instruction feedback were associated with higher intensity of movement. Greater Technical Instruction feedback was also associated with how many students were active and how often students were active during I-CAN! lessons. However, few teachers used these types of feedback to a great extent. That is, most teachers did not use positive reinforcement or technical instruction as a part of their instruction during physically active, academic lessons. Conversely, Negative feedback was inversely related to how many students active, how often students were active, and the intensity of activity. Again, few teachers utilized Negative feedback in their instruction. Consequently, these findings provide an opportunity for teacher training to enhance implementation of physically active, academic lessons. There is a great need to increase the use positive reinforcement and technical instruction. As such, emphasizing greater levels of physical activity reinforcement and technical instruction have the potential for great impact on the physical activity that is derived from these lessons. While negative feedback was not used often, lower levels of negative feedback is likely an important component to bring these outlying teachers into the greater norms of instruction.

The completion of studies 1 and 2 allowed for the development of a theoretical model with which to examine the relationships between teacher characteristics and quality of implementation as well as PA outcomes - the ultimate purpose of this dissertation. Despite limitations, these three studies, as a whole, provide insight into teacher characteristics that impact quality of implementation of physically active, academic lessons, and in turn, PA intensity. Consistent with previous work, PA-related feedback and dose of implementation were both positively associated with class PA intensity. The quality of process measure adapted in study 2 and used in study 3 focuses on the amount of PA-related feedback teachers provided students during active lessons (i.e. PA reinforcement and technical instruction). Greater amounts of PA-related feedback given during active lessons was related to higher PA intensity during lessons. A greater number of lessons implemented throughout the year was also associated with higher class PA intensity.

Study 3 also found that teacher attitudes towards implementing active lessons and PBC both positively contributed to the number of lessons implemented over the course of the year. However, only PBC was positively associated with PA-related feedback. Teacher trainings for active lessons should, therefore, focus on instructing teachers about the best types of feedback to give during lessons as well as enhancing PBC regarding implementing lessons. Additionally, attitudes seem to be less impacted by negative school climate whereas PBC is adversely impacted.

The findings from study 3 demonstrate that perceptions of less supportive, and more directive and restrictive climates are associated with less PBC. That is, teachers are less confident in their ability to implement active lessons. This is not surprising as many teachers are burdened with responsibilities related to standardized testing, which may serve as barriers to implementation. Compelling teachers to implement a program with high quality is inherently difficult. Due to high demands placed on them to get their students ready for standardized testing, implementing additional curriculum, even if only costing 10-15 minutes per day, may be very difficult to achieve. Indeed, though teachers were compensated (monetarily) for their implementation of lessons, only 70% implemented at least 4 times per week. Further, it was difficult to schedule observations of active lessons due to teachers' resistance to scheduling observation times and teacher absences. Anecdotally, teachers were resistant to scheduling observations because they perceived them as disruptions to their schedule/routine as well as taking time away from instruction. As a result, missing data led to an underpowered structural equation model.

Based on these findings, it is important to target teachers' PBC to achieve higher quality implementation. Specific barriers associated with poorer implementation quality should be assessed and direct attempts should be made to address these challenges during teacher trainings. This would provide teachers with resources and/or skills necessary to overcome barriers to implementation (that may or may not be related to demands of the school environment) and instill greater confidence in their ability to implement physically active lessons. Targeting PBC during trainings may also help teachers feel more comfortable in providing their class with PA-related feedback during active lessons. Teachers should be provided every opportunity to enhance their confidence in their ability to implement physically active lessons in order to increase program effectiveness.

While the importance of intervention implementation is generally acknowledged, it has rarely been systematically studied. The findings of this dissertation demonstrate the need for more research on this topic. School-based researchers are often dependent upon teachers to deliver interventions. Where this delivery is inconsistent the response is generally to create a new, independent variable built around the dose of implementation to capture and control for its impact. The result may be a compromise of the inherent logic of the randomized control trial to allow for an inference of causality. Hawe and colleagues (2009) suggest that fidelity of implementation should be reexamined and redefined to be less about standardization of the form of the intervention (i.e. whether the intervention looks the same across intervention sites) and instead focus on the function of the intervention (i.e. whether the intervention performs the same purpose across all sites). By redefining fidelity in this manner, this allows the intervention to adapt to initial conditions, to vary in form, rather than function.

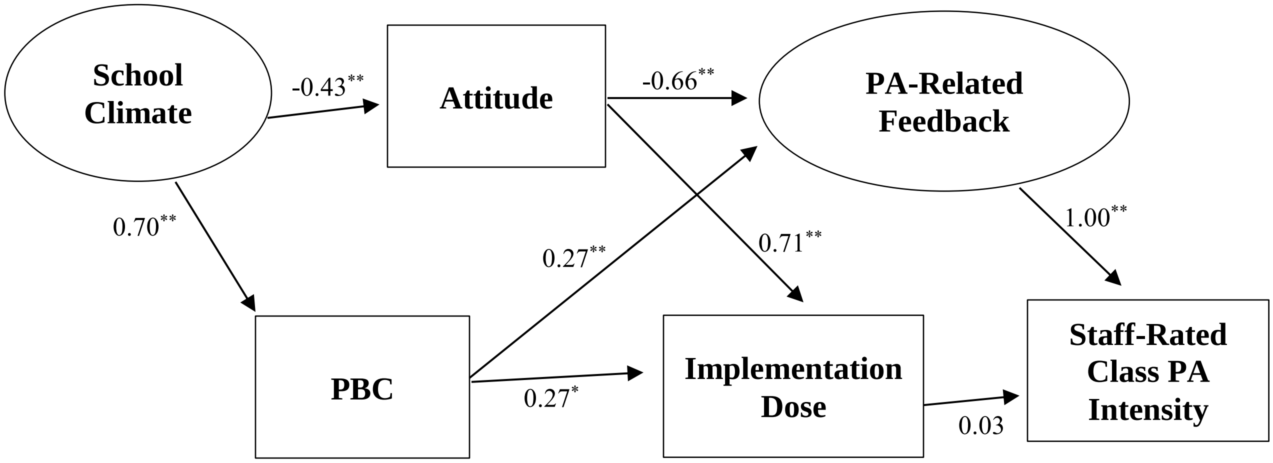
Dose of implementation is an important component of implementation quality. If teachers do not deliver the program, then students do not receive the intervention. However, implementation quality and intervention success are not just a function of dose. Implementers must be able to engage students in the intervention for students to derive maximum benefit. Indeed, it is necessary that even if students are not exposed to the recommended number of intervention days (e.g. 4 days/week), that whenever they participate in the intervention, it is a high-quality experience. Therefore, more work should be done to better understand how to train teachers for stronger, higher-quality delivery of interventions. The present data contribute to this and should help inform evaluation of similar interventions and contribute to more effective and efficient intervention development. Specifically, these studies find support for including strategies to target quality of process components such as PA-related feedback and teachers' PBC during trainings. These outcomes are not surprising given the context of the elementary classroom. Other content delivery will rely on different, local partners to deliver their content (e.g. nurses, parents, promatoras, etc.). Each of these will occur in a distinct context that is likely to be associated with a different set of predictors and needs for training. This underscores the need for more research similar to that carried out in this dissertation.

Appendix A

APRIL 2015 I-CAN LESSON / ACTIVITY EVALUATIONS
(April 6th – April 17th)

Monday	Tuesday	Wednesday	Thursday	Friday
April 6th Time: ______________ Lesson:_____________ Topic Area: _________ Location:___________ Duration Min:_______ Intensity of activity 1--2--3--4--5 Ease of Implementation 1--2--3--4--5 Student Enjoyment 1--2--3--4--5 Overall Rating 1--2--3--4--5 *Comments:*	**April 7th** Time: ______________ Lesson:_____________ Topic Area: _________ Location:___________ Duration Min:_______ Intensity of activity 1--2--3--4--5 Ease of Implementation 1--2--3--4--5 Student Enjoyment 1--2--3--4--5 Overall Rating 1--2--3--4--5 *Comments:*	**April 8th** Time: ______________ Lesson:_____________ Topic Area: _________ Location:___________ Duration Min:_______ Intensity of activity 1--2--3--4--5 Ease of Implementation 1--2--3--4--5 Student Enjoyment 1--2--3--4--5 Overall Rating 1--2--3--4--5 *Comments:*	**April 9th** Time: ______________ Lesson:_____________ Topic Area: _________ Location:___________ Duration Min:_______ Intensity of activity 1--2--3--4--5 Ease of Implementation 1--2--3--4--5 Student Enjoyment 1--2--3--4--5 Overall Rating 1--2--3--4--5 *Comments:*	Time: ______________ Lesson:_____________ Topic Area: _________ Location:___________ Duration Min:_______ Intensity of activity 1--2--3--4--5 Ease of Implementation 1--2--3--4--5 Student Enjoyment 1--2--3--4--5 Overall Rating 1--2--3--4--5 *Comments:*
April 13th Time: ______________ Lesson:_____________ Topic Area: _________ Location:___________ Duration Min:_______ Intensity of activity 1--2--3--4--5 Ease of Implementation 1--2--3--4--5 Student Enjoyment 1--2--3--4--5 Overall Rating 1--2--3--4--5 *Comments:*	**April 14th** Time: ______________ Lesson:_____________ Topic Area: _________ Location:___________ Duration Min:_______ Intensity of activity 1--2--3--4--5 Ease of Implementation 1--2--3--4--5 Student Enjoyment 1--2--3--4--5 Overall Rating 1--2--3--4--5 *Comments:*	**April 15th** Time: ______________ Lesson:_____________ Topic Area: _________ Location:___________ Duration Min:_______ Intensity of activity 1--2--3--4--5 Ease of Implementation 1--2--3--4--5 Student Enjoyment 1--2--3--4--5 Overall Rating 1--2--3--4--5 *Comments:*	**April 16th** Time: ______________ Lesson:_____________ Topic Area: _________ Location:___________ Duration Min:_______ Intensity of activity 1--2--3--4--5 Ease of Implementation 1--2--3--4--5 Student Enjoyment 1--2--3--4--5 Overall Rating 1--2--3--4--5 *Comments:*	**April 17th** Time: ______________ Lesson:_____________ Topic Area: _________ Location:___________ Duration Min:_______ Intensity of activity 1--2--3--4--5 Ease of Implementation 1--2--3--4--5 Student Enjoyment 1--2--3--4--5 Overall Rating 1--2--3--4--5 *Comments:*

Appendix B



Note: *p < 0.05, **p < 0.01

Figure 4: The Relationship between Teacher Characteristics, PA-Related Feedback, Implementation Dose, and Staff-Rated Class PA Intensity

References

Ajzen, I. (1985). From Intentions to Actions: A Theory of Planned Behavior. In P. D. J. Kuhl & D. J. Beckmann (Eds.), *Action Control* (pp. 11–39). Springer Berlin Heidelberg. Retrieved from http://link.springer.com.ezproxy.lib.utexas.edu/chapter/10.1007/978-3-642-69746-3_2

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, *19*(6), 716–723. http://doi.org/10.1109/TAC.1974.1100705

Andersen, R. E., Crespo, C. J., Bartlett, S. J., Cheskin, L. J., & Pratt, M. (1998). Relationship of physical activity and television watching with body weight and level of fatness among children: Results from the third national health and nutrition examination survey. *JAMA*, *279*(12), 938–942. http://doi.org/10.1001/jama.279.12.938

Baker, D. W., Sudano, J. J., Albert, J. M., Borawski, E. A., & Dor, A. (2001). Lack of Health Insurance and Decline in Overall Health in Late Middle Age. *New England Journal of Medicine*, *345*(15), 1106–1112. http://doi.org/10.1056/NEJMsa002887

Bartholomew, J. B., & Jowers, E. M. (2011). Physically active academic lessons in elementary children. *Preventive Medicine*, *52*, *Supplement*, S51–S54. http://doi.org/10.1016/j.ypmed.2011.01.017

Beets, M. W., Flay, B. R., Vuchinich, S., Acock, A. C., Li, K. K., & Allred, C. (2008). School Climate and Teachers' Beliefs and Attitudes Associated with

Implementation of the Positive Action Program: A Diffusion of Innovations Model. *Prevention Science*, *9*(4), 264–275. http://doi.org/10.1007/s11121-008-0100-2

Bentler, P. M., & Chou, C. P. (1987). Practical Issues in Structural Modeling. *Sociological Methods & Research*, *16*(1), 78–117. http://doi.org/10.1177/0049124187016001004

Bollen, K. A. (1989). A New Incremental Fit Index for General Structural Equation Models. *Sociological Methods & Research*, *17*(3), 303–316. http://doi.org/10.1177/0049124189017003004

Botvin, G. J., Baker, E., Dusenbury, L., Botvin, E. M., & Diaz, T. (1995). Long-Term Follow-Up Results of a Randomized Drug Abuse Prevention Trial in a White Middle-Class Population. *JAMA Journal of the American Medical Association*, *273*(14), 1106–1112.

Calle, E. E., & Kaaks, R. (2004). Overweight, obesity and cancer: epidemiological evidence and proposed mechanisms. *Nature Reviews. Cancer*, *4*(8), 579–591.

Chow, B. C., McKenzie, T. L., & Louie, L. (2008). Children's physical activity and environmental influences during elementary school physical education. Journal of Teaching in Physical Education, 27(1), 38.

Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment, 6*(4), 284.

Cothran, D. J., Kulinna, P. H., & Garn, A. C. (2010). Classroom teachers and physical activity integration. *Teaching and Teacher Education*, *26*(7), 1381–1388. http://doi.org/10.1016/j.tate.2010.04.003

Crouter, S. E., Clowers, K. G., & Bassett, D. R. (2006). A novel method for using accelerometer data to predict energy expenditure. *Journal of Applied Physiology*, *100*(4), 1324–1331. http://doi.org/10.1152/japplphysiol.00818.2005

Dale, D., Corbin, C. B., & Dale, K. S. (2000). Restricting Opportunities to Be Active during School Time: Do Children Compensate by Increasing Physical Activity Levels after School? *Research Quarterly for Exercise and Sport*, *71*(3), 240–248. http://doi.org/10.1080/02701367.2000.10608904

Daniels, S. R., Arnett, D. K., Eckel, R. H., Gidding, S. S., Hayman, L. L., Kumanyika, S., … Williams, C. L. (2005). Overweight in Children and Adolescents Pathophysiology, Consequences, Prevention, and Treatment. *Circulation*, *111*(15), 1999–2012. http://doi.org/10.1161/01.CIR.0000161369.71722.10

Datar, A., & Sturm, R. (2004). Physical Education in Elementary School and Body Mass Index: Evidence from the Early Childhood Longitudinal Study. *American Journal of Public Health*, *94*(9), 1501–1506.

Davis, J., & Wilson, S. M. (2000). Principals’ Efforts to Empower Teachers: Effects on Teacher Motivation and Job Satisfaction and Stress. *The Clearing House*, *73*(6), 349–353.

Davison, K. K., & Birch, L. L. (2001). Childhood overweight: a contextual model and recommendations for future research. *Obesity Reviews : An Official Journal of the International Association for the Study of Obesity*, *2*(3), 159–171.

DeCorby, K., Halas, J., Dixon, S., Wintrup, L., & Janzen, H. (2005). Classroom Teachers and the Challenges of Delivering Quality Physical Education. *The Journal of Educational Research*, *98*(4), 208–220.

Dehghan, M., Akhtar-Danesh, N., & Merchant, A. T. (2005). Childhood obesity, prevalence and prevention. *Nutrition Journal*, *4*(1), 24. http://doi.org/10.1186/1475-2891-4-24

Dennison, B. A., Erb, T. A., & Jenkins, P. L. (2002). Television Viewing and Television in Bedroom Associated With Overweight Risk Among Low-Income Preschool Children. *Pediatrics*, *109*(6), 1028–1035. http://doi.org/10.1542/peds.109.6.1028

Dix, K. L., Slee, P. T., Lawson, M. J., & Keeves, J. P. (2012). Implementation quality of whole-school mental health promotion and students' academic performance. *Child and Adolescent Mental Health*, *17*(1), 45–51. http://doi.org/10.1111/j.1475-3588.2011.00608.x

Domitrovich, C. E., Bradshaw, C. P., Poduska, J. M., Hoagwood, K., Buckley, J. A., Olin, S., … Ialongo, N. S. (2008). Maximizing the Implementation Quality of Evidence-Based Preventive Interventions in Schools: A Conceptual Framework. *Advances in School Mental Health Promotion*, *1*(3), 6–28. http://doi.org/10.1080/1754730X.2008.9715730

Donnelly, J. E., Greene, J. L., Gibson, C. A., Smith, B. K., Washburn, R. A., Sullivan, D. K., … Williams, S. L. (2009). Physical Activity Across the Curriculum (PAAC): A randomized controlled trial to promote physical activity and diminish overweight and obesity in elementary school children. *Preventive Medicine*, *49*(4), 336–341. http://doi.org/10.1016/j.ypmed.2009.07.022

Donnelly, J. E., & Lambourne, K. (2011). Classroom-based physical activity, cognition, and academic achievement. *Preventive Medicine*, *52*, *Supplement*, S36–S42. http://doi.org/10.1016/j.ypmed.2011.01.021

DuBose, K. D., Mayo, M. S., Gibson, C. A., Green, J. L., Hill, J. O., Jacobsen, D. J., … Donnelly, J. E. (2008). Physical activity across the curriculum (PAAC): Rationale and design. *Contemporary Clinical Trials*, *29*(1), 83–93. http://doi.org/10.1016/j.cct.2007.05.004

Dunton, G. F., Kaplan, J., Wolch, J., Jerrett, M., & Reynolds, K. D. (2009). Physical environmental correlates of childhood obesity: a systematic review. *Obesity Reviews*, *10*(4), 393–402. http://doi.org/10.1111/j.1467-789X.2009.00572.x

Durlak, J. A. (2015). Studying Program Implementation Is Not Easy but It Is Essential. *Prevention Science*, *16*(8), 1123–1127. http://doi.org/10.1007/s11121-015-0606-3

Durlak, J. A., & DuPre, E. P. (2008). Implementation Matters: A Review of Research on the Influence of Implementation on Program Outcomes and the Factors Affecting Implementation. *American Journal of Community Psychology*, *41*(3–4), 327–350. http://doi.org/10.1007/s10464-008-9165-0

Dusenbury, L., Brannigan, R., Falco, M., & Hansen, W. B. (2003). A review of research on fidelity of implementation: implications for drug abuse prevention in school settings. *Health Education Research*, *18*(2), 237–256. http://doi.org/10.1093/her/18.2.237

Dusenbury, L., Brannigan, R., Hansen, W. B., Walsh, J., & Falco, M. (2005). Quality of implementation: developing measures crucial to understanding the diffusion of preventive interventions. *Health Education Research*, *20*(3), 308–313. http://doi.org/10.1093/her/cyg134

Evenson, K. R., Catellier, D. J., Gill, K., Ondrak, K. S., & McMurray, R. G. (2008). Calibration of two objective measures of physical activity for children. *Journal of Sports Sciences*, *26*(14), 1557–1565. http://doi.org/10.1080/02640410802334196

Fairclough, S., & Stratton, G. (2005). Improving health-enhancing physical activity in girls' physical education. *Health Education Research*, *20*(4), 448–457. http://doi.org/10.1093/her/cyg137

Fakhouri, T. I., Hughes, J. P., Brody, D. J., Kit, B. K., & Ogden, C. L. (2013). Physical activity and screen-time viewing among elementary school–aged children in the united states from 2009 to 2010. *JAMA Pediatrics*, *167*(3), 223–229. http://doi.org/10.1001/2013.jamapediatrics.122

Flay, B. R., Biglan, A., Boruch, R. F., Castro, F. G., Gottfredson, D., Kellam, S., … Ji, P. (2005). Standards of Evidence: Criteria for Efficacy, Effectiveness and Dissemination. *Prevention Science*, *6*(3), 151–175. http://doi.org/10.1007/s11121-005-5553-y

Freedman, D. S., Dietz, W. H., Srinivasan, S. R., & Berenson, G. S. (1999). The relation of overweight to cardiovascular risk factors among children and adolescents: the Bogalusa Heart Study. *Pediatrics*, *103*(6 Pt 1), 1175–1182.

Freedson, P., Pober, D., & Janz, K. F. (2005). Calibration of accelerometer output for children. *Medicine and Science in Sports and Exercise*, *37*(11 Suppl), S523-530.

Gibson, C. A., Smith, B. K., Dubose, K. D., Greene, J. L., Bailey, B. W., Williams, S. L., … Donnelly, J. E. (2008). Physical activity across the curriculum: year one process evaluation results. *The International Journal Of Behavioral Nutrition And Physical Activity*, *5*, 36.

Gidlow, C. J., Cochrane, T., Davey, R., & Smith, H. (2008). In-school and out-of-school physical activity in primary and secondary school children. *Journal of Sports Sciences*, *26*(13), 1411–1419. http://doi.org/10.1080/02640410802277445

Giles, S., Jackson-Newsom, J., Pankratz, M. M., Hansen, W. B., Ringwalt, C. L., & Dusenbury, L. (2008). Measuring Quality of Delivery in a Substance Use Prevention Program. *The Journal of Primary Prevention*, *29*(6), 489–501. http://doi.org/10.1007/s10935-008-0155-7

Gillis, L. J., Kennedy, L. C., Gillis, A. M., & Bar-Or, O. (2002). Relationship between juvenile obesity, dietary energy and fat intake and physical activity. *International Journal of Obesity*, *26*(4), 458–463. http://doi.org/10.1038/sj.ijo.0801967

Gittelsohn, J., Merkle, S., Story, M., Stone, E. J., Steckler, A., Noel, J., … Ethelbah, B. (2003). School climate and implementation of the Pathways study. *Preventive*

Medicine, *37*, *Supplement 1*, S97–S106. http://doi.org/10.1016/j.ypmed.2003.08.010

Goncy, E. A., Sutherland, K. S., Farrell, A. D., Sullivan, T. N., & Doyle, S. T. (2014). Measuring Teacher Implementation in Delivery of a Bullying Prevention Program: The Impact of Instructional and Procedural Adherence and Competence on Student Responsiveness. *Prevention Science*, *16*(3), 440–450. http://doi.org/10.1007/s11121-014-0508-9

Green, L., & Kreuter, M. W. (1999). The precede–proceed model. Health promotion planning: an educational approach. 3rd ed. Mountain View (CA): Mayfield Publishing Company, 32-43.

Greenwood, C. R., Horton, B. T., & Utley, C. A. (2002). Academic Engagement: Current Perspectives in Research and Practice. *School Psychology Review*, *31*(3), 328.

Grieco, L. A., Jowers, E. M., & Bartholomew, J. B. (2009). Physically Active Academic Lessons and Time on Task: The Moderating Effect of Body Mass Index. *Medicine & Science in Sports & Exercise*, *41*(10), 1921–1926. http://doi.org/10.1249/MSS.0b013e3181a61495

Guillaume, M., Lapidus, L., & Lambert, A. (1998). Obesity and nutrition in children. The Belgian Luxembourg Child Study IV. *European Journal of Clinical Nutrition*, *52*(5), 323–328. http://doi.org/10.1038/sj.ejcn.1600532

Hallfors, D., & Godette, D. (2002). Will the `Principles of Effectiveness' improve prevention practice? Early findings from a diffusion study. *Health Education Research*, *17*(4), 461–470. http://doi.org/10.1093/her/17.4.461

Hallgren, K. A. (2012). Computing inter-rater reliability for observational data: an overview and tutorial. *Tutorials in quantitative methods for psychology, 8*(1), 23.

Harrison, F., & Jones, A. P. (2012). A framework for understanding school based physical environmental influences on childhood obesity. *Health & Place*, *18*(3), 639–648. http://doi.org/10.1016/j.healthplace.2011.12.009

Hawe, P., Shiell, A., & Riley, T. (2009). Theorising Interventions as Events in Systems. *American Journal of Community Psychology*, *43*(3–4), 267–276. http://doi.org/10.1007/s10464-009-9229-9

Hill, A. L., Degnan, K. A., Calkins, S. D., & Keane, S. P. (2006). Profiles of externalizing behavior problems for boys and girls across preschool: The roles of emotion regulation and inattention. *Developmental Psychology*, *42*(5), 913–928. http://doi.org/10.1037/0012-1649.42.5.913

Hill, J. O., & Trowbridge, F. L. (1998). Childhood Obesity: Future Directions and Research Priorities. *Pediatrics*, *101*(3), 570–574.

Hipp, J. R., & Bauer, D. J. (2006). Local solutions in the estimation of growth mixture models. *Psychological Methods*, *11*(1), 36–53. http://doi.org/10.1037/1082-989X.11.1.36

Hofferth, S. L., & Sandberg, J. F. (2001). How American Children Spend Their Time. *Journal of Marriage and Family*, *63*(2), 295–308. http://doi.org/10.1111/j.1741-3737.2001.00295.x

Hollar, D., Messiah, S. E., Lopez-Mitnik, G., Hollar, T. L., Almon, M., & Agatston, A. S. (2010). Effect of a Two-Year Obesity Prevention Intervention on Percentile

Changes in Body Mass Index and Academic Performance in Low-Income Elementary School Children. *American Journal of Public Health*, *100*(4), 646–653. http://doi.org/10.2105/AJPH.2009.165746

Horn, T. S. (1985). Coaches' feedback and changes in children's perceptions of their physical competence. *Journal of Educational Psychology*, *77*(2), 174–186. http://doi.org/10.1037/0022-0663.77.2.174

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55. http://doi.org/10.1080/10705519909540118

Keith, T. Z. (2014). *Multiple Regression and Beyond*. *An introduction to multiple regression and structural equation modeling.* Routledge.

Kibbe, D. L., Hackett, J., Hurley, M., McFarland, A., Schubert, K. G., Schultz, A., & Harris, S. (2011). Ten Years of TAKE 10!®: Integrating physical activity with academic concepts in elementary school classrooms. *Preventive Medicine*, *52*, *Supplement*, S43–S50. http://doi.org/10.1016/j.ypmed.2011.01.025

Kim, Y., Beets, M. W., & Welk, G. J. (2012). Everything you wanted to know about selecting the "right" Actigraph accelerometer cut-points for youth, but…: A systematic review. *Journal of Science and Medicine in Sport*, *15*(4), 311–321. http://doi.org/10.1016/j.jsams.2011.12.001

Kraft, R. E. (1989). Children at Play: Behavior of Children at Recess. *Journal of Physical Education, Recreation & Dance*, *60*(4), 21–24. http://doi.org/10.1080/07303084.1989.10604403

Lei, P. W., & Wu, Q. (2007). Introduction to Structural Equation Modeling: Issues and Practical Considerations. *Educational Measurement: Issues and Practice*, *26*(3), 33–43. http://doi.org/10.1111/j.1745-3992.2007.00099.x

Little, R. J. A., & Rubin, D. B. (1989). The Analysis of Social Science Data with Missing Values. *Sociological Methods & Research*, *18*(2–3), 292–326. http://doi.org/10.1177/0049124189018002004

Liu, G. C., Wilson, J. S., Qi, R., & Ying, J. (2007). Green Neighborhoods, Food Retail and Childhood Overweight: Differences by Population Density. *American Journal of Health Promotion*, *21*(4s), 317–325. http://doi.org/10.4278/0890-1171-21.4s.317

Lloyd, L.K., Cook, C. L., & Kohl, H. W. (2005). A pilot study of teachers' acceptance of a classroom-based physical activity curriculum tool: TAKE 10!® TAHPERD J. Spring, pp. 8–11.

Lo, Y., Mendell, N. R., & Rubin, D. B. (2001). Testing the Number of Components in a Normal Mixture. *Biometrika*, *88*(3), 767–778.

Loprinzi, P. D., Lee, H., Cardinal, B. J., Crespo, C. J., Andersen, R. E., & Smit, E. (2012). The Relationship of Actigraph Accelerometer Cut-Points for Estimating Physical Activity With Selected Health Outcomes. *Research Quarterly for*

Exercise and Sport, *83*(3), 422–430. http://doi.org/10.1080/02701367.2012.10599877

Mahar, M., Rowe, D., Kenny, R., & Fesperman, D. (2003). EVALUATION OF THE TAKE 10 CLASSROOM-BASED PHYSICAL ACTIVITY... : Medicine & Science in Sports & Exercise. Retrieved February 24, 2016, from http://journals.lww.com/acsm-msse/Fulltext/2003/05001/EVALUATION_OF_THE_TAKE_10_CLASSROOM_BASED_PHYSICAL.742.aspx

Mahar, M. T. (2011). Impact of short bouts of physical activity on attention-to-task in elementary school children. *Preventive Medicine*, *52*, *Supplement*, S60–S64. http://doi.org/10.1016/j.ypmed.2011.01.026

Mahar, M. T., Murphy, S. K., Rowe, D. A., Golden, J., Shields, A. T., & Raedeke, T. D. (2006). Effects of a classroom-based program on physical activity and on-task behavior. *Medicine And Science In Sports And Exercise*, *38*(12), 2086–2094.

Maillard, G., Charles, M. A., Lafay, L., Thibult, N., Vray, M., Borys, J.-M., … Romon, M. (2000). Macronutrient energy intake and adiposity in non obese prepubertal children aged 5–11 y (the Fleurbaix Laventie Ville Santé Study). *International Journal of Obesity*, *24*(12), 1608–1617. http://doi.org/10.1038/sj.ijo.0801446

Malik, V. S., Schulze, M. B., & Hu, F. B. (2006). Intake of sugar-sweetened beverages and weight gain: a systematic review. *The American Journal of Clinical Nutrition*, *84*(2), 274–288.

Martin, J., & Kulinna, P. (2005). A Social Cognitive Perspective of Physical-Activity-Related Behavior in Physical Education. *Kinesiology, Health and Sport Studies*. Retrieved from http://digitalcommons.wayne.edu/coe_khs/15

Mattocks, C., Leary, S., Ness, A., Deere, K., Saunders, J., Tilling, K., … Riddoch, C. (2007). Calibration of an accelerometer during free-living activities in children. *International Journal of Pediatric Obesity*, *2*(4), 218–226. http://doi.org/10.1080/17477160701408809

Mauriello, L. M., Driskell, M. M. H., Sherman, K. J., Johnson, S. S., Prochaska, J. M., & Prochaska, J. O. (2006). Acceptability of a School-Based Intervention for the Prevention of Adolescent Obesity. *The Journal of School Nursing*, *22*(5), 269–277. http://doi.org/10.1177/10598405060220050501

McKenzie, T. L., Sallis, J. F., & Nader, P. R. (1991). System for observing fitness instruction time. J Teach Phys Educ, 11, 195-205.

Menschik, D., Ahmed, S., Alexander, M. H., & Blum, R. M. (2008). Adolescent physical activities as predictors of young adult weight. *Archives of Pediatrics & Adolescent Medicine*, *162*(1), 29–33. http://doi.org/10.1001/archpediatrics.2007.14

Meuleman, B., & Billiet, J. (2009). A Monte Carlo sample size study: How many countries are needed for accurate multilevel SEM?. I*n Survey Research Methods* (Vol. 3, No. 1, pp. 45-58).

Molnar, B. E., Gortmaker, S. L., Bull, F. C., & Buka, S. L. (2004). Unsafe to play? Neighborhood disorder and lack of safety predict reduced physical activity among urban children and adolescents. *American Journal Of Health Promotion: AJHP*, *18*(5), 378–386.

Moore, D. S., Solmon, M., Tuuri, G., Silverman, L., Zanovec, M., Guarino, A., … Murphy, E. P. (2007). A Comparison of Children's Physical Activity Levels During School and Out-of-School Activities: 2590: Board #137 June 2 8:00 AM - 9:30 AM. *Medicine & Science in Sports & Exercise*, *39*(5). http://doi.org/10.1249/01.mss.0000274949.30346.1e

Moore, L. L., Gao, D., Bradlee, M. L., Cupples, L. A., Sundarajan-Ramamurti, A., Proctor, M. H., … Ellison, R. C. (2003). Does early physical activity predict body fat change throughout childhood? *Preventive Medicine*, *37*(1), 10–17. http://doi.org/10.1016/S0091-7435(03)00048-3

Morgan, P. J., & Hansen, V. (2008). Classroom Teachers' Perceptions of the Impact of Barriers to Teaching Physical Education on the Quality of Physical Education Programs. *Research Quarterly for Exercise and Sport*, *79*(4), 506–516. http://doi.org/10.1080/02701367.2008.10599517

Muthén, B. O., & Muthén, L. K. (2012). Mplus 7 base program. Muthén & Muthén: Los Angeles, CA.

Murray, N. G., Garza, J. C., Diamond, P. M., Hoelscher, D. M., Kelder, S., & Ward, J. L. (2008). Fitness And Academic Achievement Among 3rd And 4th Grade Students In Texas: 932. *Medicine & Science in Sports & Exercise*, *40*(Supplement), S96. http://doi.org/10.1249/01.mss.0000321867.49584.bb

National Association for Sport and Physical Education. Moving into the future: national standards for physical education, 2nd ed. Reston, VA: National Association for Sport and Physical Education; 2004.

National Association for Sport and Physical Education. Recess in elementary schools. Reston, VA: National Association for Sport and Physical Education; 2006.

Newby, P. K. (2007). Are Dietary Intakes and Eating Behaviors Related to Childhood Obesity? A Comprehensive Review of the Evidence. *The Journal of Law, Medicine & Ethics*, *35*(1), 35–60. http://doi.org/10.1111/j.1748-720X.2007.00112.x

Nicaise, V., Bois, J. E., Fairclough, S. J., Amorose, A. J., & Cogérino, G. (2007). Girls' and boys' perceptions of physical education teachers' feedback: Effects on performance and psychological responses. *Journal of Sports Sciences*, *25*(8), 915–926. http://doi.org/10.1080/02640410600898095

O'Connor, T. M., Yang, S.-J., & Nicklas, T. A. (2006). Beverage Intake Among Preschool Children and Its Effect on Weight Status. *Pediatrics*, *118*(4), e1010–e1018. http://doi.org/10.1542/peds.2005-2348

Ogden, C., & Flegal, K. (2010). Changes in terminology for childhood overweight and obesity. *Natl Health Stat Report*, (25), 1–5.

Ogden, C. L., Carroll, M. D., Kit, B. K., & Flegal, K. M. (2012). Prevalence of obesity and trends in body mass index among us children and adolescents, 1999-2010. *JAMA*, *307*(5), 483–490. http://doi.org/10.1001/jama.2012.40

Ogden, C. L., Carroll, M. D., Kit, B. K., & Flegal, K. M. (2014). Prevalence of childhood and adult obesity in the united states, 2011-2012. *JAMA*, *311*(8), 806–814. http://doi.org/10.1001/jama.2014.732

Parcel, G. S., O'Hara-Tompkins, N. M., Harrist, R. B., Basen-Engquist, K. M., McCormick, L. K., Gottlieb, N. H., & Eriksen, M. P. (1995). Diffusion of an effective tobacco prevention program. Part II: evaluation of the adoption phase. *Health Education Research*, *10*(3), 297–307. http://doi.org/10.1093/her/10.3.297

Payne, A. A., & Eckert, R. (2009). The Relative Importance of Provider, Program, School, and Community Predictors of the Implementation Quality of School-Based Prevention Programs. *Prevention Science*, *11*(2), 126–141. http://doi.org/10.1007/s11121-009-0157-6

Pearson, L. C., & Moomaw, W. (2005). The Relationship between Teacher Autonomy and Stress, Work Satisfaction, Empowerment, and Professionalism. *Educational Research Quarterly*, *29*(1), 37–53.

Pellegrini, A. D., & Bohn, C. M. (2005). The Role of Recess in Children's Cognitive Performance and School Adjustment. *Educational Researcher*, *34*(1), 13–19. http://doi.org/10.3102/0013189X034001013

Pellegrini, A. D., & Smith, P. K. (1998). Physical Activity Play: The Nature and Function of a Neglected Aspect of Play. *Child Development*, *69*(3), 577–598. http://doi.org/10.1111/j.1467-8624.1998.tb06226.x

Pettigrew, J., Graham, J. W., Miller-Day, M., Hecht, M. L., Krieger, J. L., & Shin, Y. J. (2014). Adherence and Delivery: Implementation Quality and Program Outcomes for the Seventh-Grade keepin' it REAL Program. *Prevention Science*, *16*(1), 90–99. http://doi.org/10.1007/s11121-014-0459-1

Pinhas-Hamiel, O., & Zeitler, P. (1996). Insulin resistance, obesity, and related disorders among black adolescents. *The Journal Of Pediatrics*, *129*(3), 319–320.

Pulsford, R. M., Cortina-Borja, M., Rich, C., Kinnafick, F.-E., Dezateux, C., & Griffiths, L. J. (2011). Actigraph Accelerometer-Defined Boundaries for Sedentary Behaviour and Physical Activity Intensities in 7 Year Old Children. *PLoS ONE*, *6*(8), 1–9. http://doi.org/10.1371/journal.pone.0021822

Puyau, M. R., Adolph, A. L., Vohra, F. A., & Butte, N. F. (2002). Validation and calibration of physical activity monitors in children. *Obesity Research*, *10*(3), 150–157. http://doi.org/10.1038/oby.2002.24

Puyau, M. R., Adolph, A. L., Vohra, F. A., Zakeri, I., & Butte, N. F. (2004). Prediction of Activity Energy Expenditure Using Accelerometers in Children. [Miscellaneous Article]. *Medicine & Science in Sports & Exercise*, *36*(9), 1625–1631.

Rabe-Hesketh, S., Skrondal, A., & Zheng, X. (2007). Multilevel structural equation modeling. Handbook of Latent Variable and Related Models (pp. 209-227).

Riddoch, C. J. (2004). Physical Activity Levels and Patterns of 9- and 15-yr-Old European Children. *Medicine and Science in Sports and Exercise*, *36*(1), 86–92. http://doi.org/10.1249/01.MSS.0000106174.43932.92

Ringwalt, C. L., Ennett, S., Vincus, A., Thorne, J., Rohrbach, L. A., & Simons-Rudolph, A. (2002). The Prevalence of Effective Substance Use Prevention Curricula in

U.S. Middle Schools. *Prevention Science*, *3*(4), 257–265. http://doi.org/10.1023/A:1020872424136

Rogers, E. M. (2002). Diffusion of preventive innovations. *Addictive Behaviors*, *27*(6), 989–993. http://doi.org/10.1016/S0306-4603(02)00300-3

Rohrbach, L. A., Graham, J. W., & Hansen, W. B. (1993). Diffusion of a School-Based Substance Abuse Prevention Program: Predictors of Program Implementation. *Preventive Medicine*, *22*(2), 237–260. http://doi.org/10.1006/pmed.1993.1020

Sallis, J. F., & Glanz, K. (2006). The Role of Built Environments in Physical Activity, Eating, and Obesity in Childhood. *The Future of Children*, *16*(1), 89–108. http://doi.org/10.1353/foc.2006.0009

Sallis, J. F., McKenzie, T. L., Beets, M. W., Beighle, A., Erwin, H., & Lee, S. (2012). Physical Education's Role in Public Health. *Research Quarterly for Exercise and Sport*, *83*(2), 125–135. http://doi.org/10.1080/02701367.2012.10599842

Schwarz, G. (1978). Estimating the Dimension of a Model. *The Annals of Statistics*, *6*(2), 461–464.

Shrewsbury, V., & Wardle, J. (2008). Socioeconomic Status and Adiposity in Childhood: A Systematic Review of Cross-sectional Studies 1990–2005. *Obesity*, *16*(2), 275–284. http://doi.org/10.1038/oby.2007.35

Silverman, S., Tyson, L., & Krampitz, J. (1992). Teacher feedback and achievement in physical education: Interaction with student practice. *Teaching and Teacher Education*, *8*(4), 333–344. http://doi.org/10.1016/0742-051X(92)90060-G

Silvia, E. S., Thorne, J., & Tashjian, C. A. (1997). School-Based Drug Prevention Programs: A Longitudinal Study in Selected School Districts. Final Report. Retrieved from http://eric.ed.gov/?id=ED416433

Singh, A. S., Chinapaw, M. J. M., Brug, J., & van Mechelen, W. (2009). Process evaluation of a school-based weight gain prevention program: The Dutch Obesity Intervention in Teenagers (DOiT). *Health Education Research*, *24*(5), 772–777. http://doi.org/10.1093/her/cyp011

Singh, G. K., Siahpush, M., & Kogan, M. D. (2010). Neighborhood Socioeconomic Conditions, Built Environments, And Childhood Obesity. *Health Affairs*, *29*(3), 503–512. http://doi.org/10.1377/hlthaff.2009.0730

Smith, R. E., & Smoll, F. L. (1990). Self-esteem and children's reactions to youth sport coaching behaviors: A field study of self-enhancement processes. *Developmental Psychology*, *26*(6), 987–993. http://doi.org/10.1037/0012-1649.26.6.987

Smith, R. E., & Smoll, F. L. (1996). The coach as a focus of research and intervention in youth sports. Children and youth in sport: A biopsychosocial perspective, 125-141.

Standage, M., Duda, J. L., & Ntoumanis, N. (2003). A model of contextual motivation in physical education: Using constructs from self-determination and achievement goal theories to predict physical activity intentions. *Journal of Educational Psychology*, *95*(1), 97–110. http://doi.org/10.1037/0022-0663.95.1.97

Stewart, J. A., Dennison, D. A., Kohl, H. W., & Doyle, J. A. (2004). Exercise Level and Energy Expenditure in the TAKE 10!® In-Class Physical Activity Program.

Journal of School Health, *74*(10), 397–400. http://doi.org/10.1111/j.1746-1561.2004.tb06605.x

Stokols, D. (1996). Translating Social Ecological Theory into Guidelines for Community Health Promotion. *American Journal of Health Promotion*, *10*(4), 282–298. http://doi.org/10.4278/0890-1171-10.4.282

Story, M., Nanney, M. S., & Schwartz, M. B. (2009). Schools and Obesity Prevention: Creating School Environments and Policies to Promote Healthy Eating and Physical Activity. *The Milbank Quarterly*, *87*(1), 71–100. http://doi.org/10.1111/j.1468-0009.2009.00548.x

Tein, J. Y., Coxe, S., & Cham, H. (2013). Statistical Power to Detect the Correct Number of Classes in Latent Profile Analysis. *Structural Equation Modeling: A Multidisciplinary Journal*, *20*(4), 640–657. http://doi.org/10.1080/10705511.2013.824781

Treuth, M. S., Schmitz, K., Catellier, D. J., Mcmurray, R. G., Murray, D. M., Almeida, M. J., … Pate, R. (2004). Defining Accelerometer Thresholds for Activity Intensities in Adolescent Girls. *Medicine and Science in Sports and Exercise*, *36*(7), 1259.

Trickett, E. J., & Birman, D. (1989). Taking ecology seriously: A community development approach to individually based preventive interventions in schools. In L. A. Bond & B. E. Compas (Eds.), *Primary prevention and promotion in the schools* (pp. 361–390). Thousand Oaks, CA, US: Sage Publications, Inc.

Trost, S. G., Kerr, L. M., Ward, D. S., & Pate, R. R. (2001). Physical activity and determinants of physical activity in obese and non-obese children. *International Journal of Obesity*, *25*(6), 822–829. http://doi.org/10.1038/sj.ijo.0801621

Trost, S. G., Loprinzi, P. D., Moore, R., & Pfeiffer, K. A. (2011). Comparison of Accelerometer Cut Points for Predicting Activity Intensity in Youth: *Medicine & Science in Sports & Exercise*, *43*(7), 1360–1368. http://doi.org/10.1249/MSS.0b013e318206476e

Tucker, L. A., Seljaas, G. T., & Hager, R. L. (1997). Body Fat Percentage of Children Varies According to Their Diet Composition. *Journal of the American Dietetic Association*, *97*(9), 981–986. http://doi.org/10.1016/S0002-8223(97)00237-X

Wang, Y. (2001). Cross-national comparison of childhood obesity: the epidemic and the relationship between obesity and socioeconomic status. *International Journal of Epidemiology*, *30*(5), 1129–1136. http://doi.org/10.1093/ije/30.5.1129

Weiss, M. R. (2000). Motivating Kids in Physical Activity. *President's Council on Physical Fitness and Sports Research Digest*. Retrieved from http://eric.ed.gov/?id=ED470695

Whitaker, R. C., Wright, J. A., Pepe, M. S., Seidel, K. D., & Dietz, W. H. (1997). Predicting Obesity in Young Adulthood from Childhood and Parental Obesity. *New England Journal of Medicine*, *337*(13), 869–873. http://doi.org/10.1056/NEJM199709253371301

Wieting, J. (2008). Cause and effect in childhood obesity: solutions for a national epidemic. *The Journal of the American Osteopathic Association*, *108*(10), 545–552.

Williams, S., Kibbe, D.L., & Lombardo, M.A. (2008). Delta H.O.P.E. Tri-State Initiative, August 2003 to June 2007: final report to the Mississippi Alliance for Self-Sufficiency and the W.K. Kellogg Foundation.

Wilson, L. F. (2007). Adolescents' Attitudes About Obesity and What They Want in Obesity Prevention Programs. *The Journal of School Nursing*, *23*(4), 229–238. http://doi.org/10.1177/10598405070230040801

Wolf, E. J., Harrington, K. M., Clark, S. L., & Miller, M. W. (2013). Sample Size Requirements for Structural Equation Models An Evaluation of Power, Bias, and Solution Propriety. *Educational and Psychological Measurement*, *73*(6), 913–934. http://doi.org/10.1177/0013164413495237

Wolfinger, R., & Chang, M. (1995). Proceedings of the Twentieth Annual SAS Users Group Conference.

Wurpts, I. C., & Geiser, C. (2014). Is adding more indicators to a latent class analysis beneficial or detrimental? Results of a Monte-Carlo study. *Frontiers in Psychology*, *5*. http://doi.org/10.3389/fpsyg.2014.00920

Xiang, P., Lowy, S., & McBride, R. (2002). The Impact of a Field-Based Elementary Physical Education Methods Course on Preservice Classroom Teachers' Beliefs. *Journal of Teaching in Physical Education*, *21*(2), 145–61.