2015

Two Studies To Inform Comprehensive School Physical Activity Programming: A Systematic Review of Program Effectiveness and the Development of an Observational Measure for Classroom-Based Physical Activity Promotion

Laura B. Russ
University of South Carolina

TWO STUDIES TO INFORM COMPREHENSIVE SCHOOL PHYSICAL ACTIVITY PROGRAMMING: A SYSTEMATIC REVIEW OF PROGRAM EFFECTIVENESS AND THE DEVELOPMENT OF AN OBSERVATIONAL MEASURE FOR CLASSROOM-BASED PHYSICAL ACTIVITY PROMOTION

by

Laura B. Russ

Bachelor of Science
University of Georgia, 2003

Master of Arts
Augusta State University, 2007

Submitted in Partial Fulfillment of the Requirements

For the Degree of Doctor of Philosophy in

Physical Education

College of Education

University of South Carolina

2015

Accepted by:

Collin Webster, Major Professor

Michael Beets, Committee Member

Rachel Harvey, Committee Member

David Phillips, Committee Member

Lacy Ford, Senior Vice Provost and Dean of Graduate Studies

Dedication

To God be the glory: That others may see and also believe (Psalm 40:2-4).

To Stephen: Thank you for your belief, your continued support, and your encouragement. You always knew when I needed it most, and I would not be where I am today without you. Thank you for running alongside me during yet another marathon in our life together. I will always thank you for this opportunity and your selflessness and steadfastness throughout the process. You keep me grounded, give me perspective, and carry me forward when I don’t have the strength to walk. You know the rest.

Acknowledgements

Dr. Webster: You have believed in me, challenged me, and refined me. I recognize and appreciate your investment in my educational growth and patience during the process. I will forever embrace thankfulness for having you as a mentor. Thank you is simply not enough.

Dr. Doutis: Thank you for inspiring me, teaching me, and challenging me. I am a better person for knowing you and will always have the highest respect for your level of dedication, investment, and professionalism.

Dr. Beets: Thank you for the opportunities to learn, providing me with insight, feedback, and important questions to consider. You have helped me develop the ability to look beyond the obvious.

Dr. Harvey: Thank you for your continued support and feedback. You taught me how to give myself grace.

Dr. Phillips: Thank you for appreciating me as a person, and reminding me that I am more than my work. That has been very valuable to me.

Abstract

This dissertation consists of two studies pertaining to comprehensive school physical activity programs (CSPAP) that function in tandem to advance the knowledge base. The lack of an empirical basis for moving forward with CSPAP efforts and the lack of objective measures of CSPAP implementation are intertwined limitations currently stemming the potential for wide scale program adoption.

The purpose of Study 1 was to conduct a systematic review and meta-analysis of multi-component PA interventions through schools that could be mapped onto at least two components of the CSPAP model. Electronic databases were searched to identify published studies that (1) occurred in the US; (2) targeted K-12; (3) were interventions; (4) reflected ≥2 CSPAP components, with at least one targeting school-based PA during school hours; and (5) reported outcomes as improvements in daily PA. Standardized mean effects (Hedge's *g*) from pooled random effects inverse-variance models were estimated. The overall impact of interventions was small (0.11, 95CI 0.03 to 0.19) with more CSPAP components related to increased effectiveness (effect size of 0.06, 0.19, and 0.29 corresponding with 2, 3, and 4 components, respectively). Studies employing objective measures of PA (n=3) resulted in smaller effects (0.02 vs. 0.12) than those using self-report (n=14). Studies including PADSD (0.19 vs. 0.07) and SW (0.21 vs. 0.09) were associated with a larger effect size than interventions not including these components. As designed, there is limited evidence of the effectiveness of multi-

component interventions to increase youth total daily PA. Results suggest that taking a multi-component approach to increasing youth PA is an appropriate path, but strategies within and across components may need to be reconsidered for maximal impact.

The purpose of Study 2 was to describe instrument development, reliability, and validity of the System for Observing Student Movement during Academic Routines and Transitions (SOSMART). An extensive literature review and Delphi survey were used in developing an *a priori* framework to guide live observations of purposefully selected classroom teachers. Examples of movement integration (MI) were considered in light of the initial framework and expanded and/or refined as needed. Reliability was tested using intra and interobserver percent agreement. Two validity procedures were used in this study. The Delphi survey was used to further examine content validity, and multilevel random effects logistical regression models were estimated for each of the MI variables to test construct validity of the instrument by examining the presence/absence of teacher MI compared with students' activity and/or sedentary behaviors as measured with accelerometers.

Intraobserver agreement across two weeks resulted in 97.5% agreement and interobserver agreement exceeded 80% in live and video reliability testing. Results support the hypothesis that a student was more likely to be in activity when MI variables were present in the same minute with 8 out of 11 variables achieving statistical significance. Three MI variables were not sufficiently observed (i.e. reward, other movement (academic), physical environment); therefore, reliability and construct validity was not calculated for these variables. Continued use of SOSMART is needed to further validate these variables. Future research utilizing SOSMART can provide descriptive information about the extent of MI in classrooms, which MI strategies may be more or less effective in certain contexts, and explore reasons for any differences in activity outcomes as a result of MI. This information can also be used to create a national benchmark for MI in the classroom and potentially influence the practice of teacher evaluations by administrators.

Together, these studies contribute to the foundational knowledge for CSPAP research and have potential to impact policy and practice decisions in pre-service teacher education, in-service teacher development, and future PA research.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

This dissertation consists of two studies pertaining to Comprehensive School Physical Activity Programs (CSPAP). The purpose of this chapter is to provide the reader with a brief overview of the need for these studies and how they can function in tandem to advance the knowledge base. Specifically, the chapter identifies the lack of an empirical basis for moving forward with CSPAP efforts and the lack of objective measures of CSPAP implementation as intertwined limitations currently stemming the potential for wide scale program adoption. The chapter concludes with the purpose of each study.

Background

There exists a plethora of research illustrating the importance of PA for children (Centers for Disease Control and Prevention [CDC], 2013; CDC, 2010; Institute of Medicine [IOM], 2013). Benefits of PA range from decreasing anxiety and/or depression and the level of physical health risk factors (i.e. Type 2 diabetes) to increasing self-esteem, academic performance, and physical health performance (i.e. muscle and bone strength) (CDC, 2010; CDC, 2013; McKenzie & Kahan, 2008), providing evidence that PA is important to the physical and mental health of children (IOM, 2013). Unfortunately, America’s youth are not meeting the national recommendation for 60 minutes or more of moderate- to vigorous-PA each day (CDC, 2013; United States

Department of Health and Human Services [USDHHS], 2008). There is growing concern that children are becoming more sedentary, thereby indicating that reducing sedentary time may be just as important as efforts to increase PA (IOM, 2013).

Since children are in schools for the majority of their waking hours during the week (CDC, 2013; IOM, 2013; National Physical Activity Plan, 2012), it is no surprise that schools have been identified as a key setting to intervene. Recommendations for a "whole-of-school" approach include implementing CSPAPs (CDC, 2013; IOM, 2013). Unfortunately, school-based efforts, when considered in their entirety, have been minimally effective (Russ, L., Webster, C., Beets, M., & Phillips, D., 2015). A lack of empirical evidence of CSPAP effectiveness has hampered progress of program adoption. Moreover, CSPAP efforts have lacked objective measures of implementation, thereby not providing empirical evidence to advance the knowledge base informing CSPAP efforts.

Limited Empirical Basis for CSPAP

While many interventions through schools target youth PA with minimal impact (Metcalf, Wilkin, & Henry, 2012; van Sluijs, McMinn, & Griffin, 2007), little is known about the effectiveness of multi-component PA interventions through schools, reflecting the recommended whole-of-school approach. For example, in a review of 33 controlled trials targeting children's PA, 10 studies were categorized as multi-component and collectively yielded inconclusive results of effectiveness (van Sluijs et al., 2007). A more recent review of children's PA intervention effectiveness examined 30 studies that were school-based or home/family based but did not pay specific attention to multiple components of an intervention working synergistically (Metcalf et al., 2012). No empirical evidence exists documenting the effectiveness of multi-component PA interventions through schools in alignment with CSPAPs; yet implementing CSPAPs is presently recommended by leading national organizations (CDC, 2013; IOM, 2013; National Association of Sport and Physical Education [NASPE], 2013).

Lack of Objective Measures of CSPAP Implementation

At the state and national level, teachers have provided survey responses about the extent to which they are providing opportunities for students to be physically active (American Alliance for Health, Physical Education, Recreation, and Dance [AAHPERD], 2011; Elmakis, 2010). Despite using multiple data sources, none of them included objective measures of implementation. Empirical evidence documenting the effectiveness of CSPAPs is needed to provide support for pursuing this approach as a viable path for impacting youth PA. However, there is a lack of objective measurements within and across CSPAP components to provide such evidence. In order to describe and evaluate the impact of CSPAPs, component-specific objective measures of implementation are needed. One such component is PA during the school day.

Providing opportunities for students to be active during the school day places classroom teachers in the spotlight because they have students in their care for the majority of the school day. In addition, PA can occur in a variety of settings during the school day, including lunch, recess, and the academic classroom-all of which involve the classroom teacher. Thus, it is not surprising that movement integration (MI) is a strategy recommended to classroom teachers for helping students accrue minutes of activity

(IOM, 2013; Webster, Russ, Vazou, Goh, & Erwin, 2015). Unfortunately, the extent to which teachers are implementing movement integration (MI) is limited to self-reports (Bartholomew & Jowers, 2011; Cradock et al., 2014; Gibson et al., 2008; Howie, Newman-Norland, & Pate, 2014; Kohl, Moore, Sutton, Kibbe, & Schneider, 2001; Kibbe et al., 2011; Skrade, 2013; Stewart, Dennison, Kohl, & Doyle, 2004; Webster et al., 2013; Williamson et al., 2007; Woods, 2011). At least part of the explanation for such reliance on self-reports can be explained by the lack of an objective measurement tool designed to capture MI.

The limited presence of evidence is entangled with the issue of lacking objective measurements capable of providing such data. The related nature of these two issues means progress toward one (e.g. providing empirical evidence) is thwarted until advancements are made toward the other (e.g. creating objective instrumentation to measure CSPAP implementation). Empirical evidence cannot be provided without objective measurement tools. Research providing empirical evidence of CSPAP outcomes can strengthen the perspective that CSPAPs are an effective avenue to helping students meet PA recommendations; thus, creating a justification for continuing to pursue these programs. However, trying to document such evidence solely with self-reported data is not sufficient. Objective measures of component-specific implementation will provide stronger evidence about the strengths and limitations of each component. Together, these studies serve to contribute to the knowledge base serving as a foundation for CSPAP research and have potential to impact policy and practice decisions in preservice teacher education, inservice teacher development, and future PA research.

Therefore, the following studies function in tandem to address these two issues thereby collectively advancing the knowledge base needed for evaluating CSPAP effectiveness.

Purpose of the Studies

The purpose of Study 1 was to conduct a systematic review and meta-analysis of multi-component PA interventions through schools that could be mapped onto at least two components of the CSPAP model. Specific research questions this study addressed were:

- To what extent are multi-component PA interventions through schools effective?
- To what extent does the effect vary across the number of CSPAP components targeted?
- Which CSPAP components are associated with a greater effect size?

The purpose of Study 2 was to describe the instrument development, reliability, and validity of a System for Observing Student Movement During Academic Routines and Transitions (SOMART). Specific research questions this study addressed were:

- What types of physical activity promotion strategies are being utilized by teachers in elementary general education classrooms?
- What coding scheme can be developed to measure the items above?
- To what extent is SOSMART a valid measure of physical activity promotion and able to be used reliably?

These two studies are related in that they each address a limitation currently stunting the progress and impact of CSPAP research.

CHAPTER 2

LITERATURE REVIEW

The purpose of this chapter is to provide the reader with an extended perspective of the literature informing the second study presented in this dissertation. A review of literature for the first study is excluded from this chapter, as the first study is itself a systematic review. The chapter is organized into the following sections: (a) importance of increasing PA and decreasing sedentarism in children, (b) schools as an important setting for intervention, (c) classroom movement integration as a key recommended strategy, and (d) the need for objective measures of classroom movement integration.

Importance of Increasing Physical Activity and Decreasing Sedentarism in Children

Physical activity (PA) is well documented as important and beneficial for children in many ways (CDC, 2013; CDC, 2010; IOM, 2013). Increasing PA is associated with improved health through reducing risk factors for diseases like obesity, Type 2 diabetes, and cardiovascular disease (CDC, 2013; McKenzie & Kahan, 2008). Being active is also associated with improvements to muscular strength, bone strength, self-esteem, and lower levels of anxiety and/or depression (CDC, 2013), thereby demonstrating the importance of PA to the mental and physical health of children (IOM, 2013). Further, increased amounts of PA during school have been associated with improved academic performance of children (CDC, 2010). For example, an extensive review of literature conducted by the CDC (2010) found the majority of studies on PA provided evidence in support of a positive association between PA and academic achievement, skills, and behaviors. In addition to the potential of improving student academic performance, the report confirmed that increasing opportunities for students to be physically active in schools will not result in declining academic performance.

The U.S. Department of Health and Human Services (2008) recommends America's youth (6+ years old) engage in 60 minutes or more of moderate- to vigorous-intensity PA every day. Not only are children not meeting this recommendation (CDC, 2013; USDHHS, 2008), but they are also engaging in sedentary behaviors in their classrooms, which is where they spend up to 9 hours each school day (CDC, 2013). While increasing opportunities for PA is important, reducing sedentary time may be equally important (IOM, 2013).

Due to the variability in definitions for sedentarism, the Institute of Medicine (2013) describes sedentary behavior in terms of what it is not. Sedentary behavior is when an individual is not engaged in sleeping, light-, moderate-, or vigorous-intensity activity. Sedentary behaviors may further be classified into two categories: recreational (e.g., "screen time" such as watching television or reading for pleasure) or non-recreational (e.g., schoolwork or other sedentary daily tasks such as eating or driving to work). The concern that children are increasingly sedentary (IOM, 2013) is reflected in research that has used interventions designed to target and reduce sedentary behaviors of children and youth (Gortmaker et al., 1999; Robinson, 1999; Salmon et al., 2005). With evidence suggesting children's health may be negatively affected through accumulated sedentary behaviors despite their engagement in PA (Biddle, Gorley, & Stensel, 2004;

Dietz, 2001; Salmon, 2010), it is important to decrease the amount of sedentary opportunities children have during the day.

Schools as an Important Setting for Intervention

One approach is to find ways to change non-recreational sedentarism during the school day. Many interventions targeting increases in PA or decreases in sedentary behaviors of children have taken place in schools (Russ, Webster, Beets, & Phillips, 2015; van Sluijs et al., 2007). Schools are a unique and promising setting to help children meet PA guidelines (CDC, 2013; IOM, 2013; National Physical Activity Plan, 2012; Pate et al., 2006; USDHHS, 2012) and have historically played a role in children's health. Schools have access to most children regardless of race, ethnicity, or socioeconomic status for most waking hours on weekdays, and can improve academic performance through PA (CDC, 2013; IOM, 2013; National Physical Activity Plan, 2012) making them a unique setting to target children's PA levels. Recommendations for helping children increase PA and decrease non-recreational sedentarism include utilizing a school-wide multi-component approach (CDC, 2013; IOM, 2013).

This "whole-of-school" approach is defined as "all of a school's components and resources operat[ing] in a coordinated and dynamic manner to provide access, encouragement, and programs that enable all students to engage in vigorous- or moderate-intensity physical activity 60 minutes or more each day" (IOM, 2013, p. 367). An example of this coordinated approach is a Comprehensive School Physical Activity Program (CSPAP). Distinct components of a CSPAP include 1) quality physical education, 2) PA during the school day, 3) PA before or after school, 4) staff involvement, and 5) family and community engagement (CDC, 2013). With access to children, facilities, equipment, and staff, schools have the foundation already in place to facilitate a school-wide approach to PA promotion. Unfortunately, an overwhelming amount of schools do not provide students with enough opportunities to be physically active (CDC, 2013; Lee, Burgeson, Fulton, & Spain, 2007). Moreover, as students spend greater amounts of time away from home and in schools, it is increasingly urgent to maximize the potential of each school component to create opportunities for students to be active (Sturm, 2005).

Classroom Movement Integration as a Key Recommended Strategy

The academic classroom is a setting with potential to integrate movement opportunities for students because elementary students spend most of their day in the academic classroom with their teacher (Kohl et al., 2001; IOM, 2013). This means classroom teachers have access to students during the school day that other faculty and staff (i.e., PE teachers) do not have. Thus, the academic classroom is a setting in schools with potential to be an effective component in a whole-of-school approach to PA (Pangrazi, Beighle, Vehige, & Vack, …2003; Stewart et al., 2004) and help students meet PA recommendations (Bartholomew & Jowers, 2011).

Classroom-based PA "includes all activity regardless of intensity performed in the classroom during normal classroom time" (IOM, 2013, p. 266). This includes movement integration during academic lessons, movement used as breaks between lessons, and even movement in special area subjects (e.g., Art). This definition does not include activity during physical education (PE), recess, or lunch breaks. Not only is there empirical evidence showing classroom-based PA can facilitate contributions to student PA

(Bartholomew & Jowers, 2011; Beighle, Erwin, Beets, Morgan, & Le Masurier, 2010; Erwin, Beighle, Morgan, & Noland, 2011; Holt, Bartee, & Heelan, 2012; Mahar et al., 2006), but also that these contributions account for up to 19 minutes of the national recommendation for 60 minutes or more of moderate- to vigorous-intensity PA (Bassett et al., 2013). Other positive results from classroom-based PA include decreasing sedentarism (Gortmaker et. al, 1999; Robinson, 1999; Salmon et al., 2005; Salmon, 2010), improving on-task behavior (Grieco, Jowers, & Bartholomew, 2009; Howie, 2013; Mahar et al., 2006; Mahar, 2011), positive affect (Howie et al., 2014), and cognitive function (Donnelly & Lambourne, 2011; Elmakis, 2010; Howie et al., 2014). With increasing pressure for performance on high stakes testing, classroom-based PA offers teachers a way to enhance student achievement, contribute to meeting national PA recommendations, and reduce non-recreational sedentarism without compromising academic performance (CDC, 2013; IOM, 2013). Despite these benefits, little is known about the extent to which classroom-based PA, or movement integration, occurs.

The Need for Objective Measures of Classroom Movement Integration

Objective measurements of movement integration in the classroom are needed for several reasons. First, objective measurements can provide empirical evidence of the frequency and variety of MI, thus contributing to a descriptive knowledge base. Second, objective measurements can document fidelity of implementation in intervention settings. Third, objective measurements can create a common way to communicate about MI, through creating and utilizing working definitions for MI behaviors that can be used to prepare preservice teachers, inservice professionals, and inform future research using a whole-of-school approach to increasing students' PA and/or decreasing sedentarism. The following section provides support for the need to have objective measurements of classroom movement integration and is organized into the following subsections: a) descriptive-analytic research, b) implementation fidelity, c) variables related to MI implementation, and d) contributions of systematic observation.

Descriptive-analytic data. An extensive literature search yielded a paucity of research on MI. There is a lack of descriptive research on MI in these settings resulting in limited knowledge about what transpires in the academic classroom. That there is a wide-scale problem with the current status of opportunities to be active or sedentary in the academic classroom is simply not documented.

Only two surveys provided any descriptive information about the nature and/or extent of MI in the absence of policy (Elmakis, 2010; AAHPERD, 2011). A graduate student at the College of William and Mary surveyed CTs across the state of Virginia to find out the extent to which PA was incorporated into classrooms (Elmakis, 2010). Using an unpublished survey instrument developed for the study (Physical Activity in the Classroom), 393 elementary school teachers responded to questions asking how many minutes they spent devoted to PA in their academic lessons (outside of recess and physical education), which content areas they used the most to incorporate PA, and if they were likely to incorporate more PA during the school day. Results indicated low levels of PA incorporated into academic lessons with math and science as the academic content used most often. Despite the small extent to which teachers self-reported MI, results also indicated teachers expressed willingness to do more if provided with various supports.

In another survey, the American Alliance for Health, Physical Education, Recreation, and Dance (AAHPERD) canvassed the nation to better understand the extent to which US schools are providing PA opportunities for students aligned with the CSPAP model (AAHPERD, 2011). The baseline data from the results indicated around half of the elementary schools integrated PA between lessons, and less than half of the elementary schools reported promoting PA within academic lessons or at the beginning of the school day.

These studies provide limited information about the extent of MI, do not provide a clear and objective picture of MI, and do not provide substantial documentation of what is taking place in the academic classroom. While the knowledge gleaned about the extent and nature of MI from research is scarce, it is also reliant on self-reported data.

Implementation fidelity. Classroom-based PA may not always be implemented as designed. Measuring implementation of interventions is important because it permits progress monitoring and identifies areas in need of revision or removal (McGraw et al., 2000). McGraw and colleagues (2000) conceptualize implementation measurements of programs and policies promoting PA as either quantitative (reflecting completeness) or qualitative (fidelity), and describe how teacher self-reports may result in overestimating actual implementation rates. Since self-reported teacher implementation rates may impact the calculated effectiveness of classroom-based PA interventions (Bartholomew & Jowers, 2011; Donnelly et al., 2009), it is critical to have accurate information on the fidelity of MI implementation.

Existing data on fidelity of MI in intervention settings has relied almost solely on self-report. For example, in a review of Take10! interventions that occurred over a 10 year period, including 19 instances of implementation, not a single objective measurement was utilized to evaluate implementation fidelity (Kibbe et al., 2011). Commonly used self-reports included weekly PA logs (Cradock et al., 2014; Naylor, Macdonld, Zebedee, Reed, & McKay, 2006; Skrade, 2013; Stewart et al., 2004; Woods, 2011), teacher surveys (Cradock et al., 2014; Dubose et al., 2008; Holt et al., 2013; Naylor et al., 2006; Williamson, 2007), and teacher focus groups (Gibson et al., 2008; Howie et al., 2014; Naylor et al., 2006). Even recommendations for monitoring intervention implementation fidelity have centered on teacher and student self-reports (i.e., completing daily logs on activity type and duration) (Erwin et al., 2011). Self-reports do provide us with one perspective; however, they typically do not provide the most accurate data. When compared to direct measures of PA, self-report measures can result in both overestimations and underestimations of PA (Prince et al., 2008).

Variables related to MI implementation. In their mediating variable framework, Baranowski and Jago (2005) suggest there are many factors that can impact a teacher's implementation of a new program. At any number of points in time during the implementation processes, teachers face different barriers to implementation. For example, Gibson and colleagues (2008) were able to identify barriers to implementation reported by the teachers (e.g., needing lessons that could be used in small classrooms, less "babyish" lessons, and time constraints) which suggests that not monitoring implementation could have resulted in different effects due to the impact of barriers and other factors that can affect implementation. Thus, monitoring implementation and giving consideration to variables related to MI implementation can identify potential areas of weakness and is therefore critical.

Teacher training. If CTs are to be expected to implement MI strategies, in intervention or non-intervention settings, they must receive training on the methods and procedures required. The most common type of training documented in the literature was in-service or professional development days (Bartholomew & Jowers, 2011; Cradock et al., 2014; DuBose et al., 2008; Dunn, Venturanza, Walsh, & Nonas, 2010; Erwin et al., 2011; Holt et al., 2013; Mahar, et al., 2006; Naylor et al., 2006; Stewart et al., 2004; Woods, 2011) ranging in duration from 30 minute sessions to all day. A few studies offered ongoing training (Cothran, Kulinna, & Garn, 2010; DuBose et al., 2008), and two different studies offered additional support via telephone or booster sessions (Naylor et al., 2006; Woods, 2011). There were two cases where CTs were not explicitly trained. For example, Skrade (2013) describes orienting CTs to Move-For-Thought (M4T) by simply giving the materials to teachers to take home and read. The other case involves incidental MI, PA not explicitly directed by the teacher, where CTs were encouraged to use the exercise balls themselves in addition to receiving a resource booklet with activities incorporating the exercise balls; however, no explicit training was required because the nature of the intervention, by design, was not teacher directed (Janulewicz, 2008).

The training sessions varied in the presence of hands-on activities or participation (Bartholomew & Jowers, 2011; Cothran et al., 2010; Dunn et al., 2010; Erwin et al., 2011; Holt et al., 2013; Mahar et al., 2006; Stewart et al., 2004; Woods, 2011), tangible resources provided (e.g., equipment, lesson plans, or activity books) (Bartholomew &

Jowers, 2011; Cradock et al., 2014; DuBose et al., 2008; Dunn et al., 2010; Holt et al., 2013; Mahar et al., 2006; Naylor et al., 2006; Stewart et al., 2004), and opportunities for collaboration (Bartholomew & Jowers, 2011; Cothran et al., 2010; Dunn et al., 2010; Erwin et al., 2011); however, the main focus was typically on deliberate MI through exercise breaks or integrating movement into academic content. No trainings described any strategies focused on incidental MI.

Policy. Promotion of PA can be targeted through policy mandates. As such, the presence or absence of policy may be a factor linked to rates of implementation. Are teachers more likely to integrate movement if there is a policy in place? Three studies examined the extent to which CTs adopted or implemented MI in relation to state or district policies (Evenson, Ballard, Lee, & Ammerman, 2009; Holt et al., 2013; Webster et al., 2013). In North Carolina, of 106 responding school districts, 45% of elementary schools reported using a pre-packaged program (i.e. 34% used Energizers and less than 11% used Take 10!) for classroom-based PA (Evenson et al., 2009). Teachers from four elementary schools in a rural district in Nebraska reported the number of days they met the mandate of 20 minutes of PA daily. Over the course of the academic year, teachers promoting PA declined from 40% of teachers reporting they met the policy requirement in September to only 4% in February of the same academic year (Holt et al., 2013). Finally, in South Carolina, 201 elementary CTs were surveyed about MI through a six item questionnaire assessing the frequency of PA promotion behaviors aligned with current recommendations (Webster et al., 2013). Results revealed a mean score of 2.11 on a 5-point scale (0=Never, 5=Very Often) suggesting elementary classroom teachers only “sometimes” promoted PA in the classroom.

Physical space. Barriers such as physical space constraints or large class sizes (Gibson et al., 2008; Goh et al., 2013) are contextual factors that may influence MI implementation and are likely to be directly observed. Therefore, physical space constraints may warrant consideration from an observation instrument development perspective. Overall, the knowledge base does not document many facilitators and/or barriers to MI that are directly observable.

We not only have little knowledge about MI in the literature, but we also have little understanding of what observable facilitators and/or barriers exist to MI, as well as how they function. Part of the reason we have limited knowledge about the extent to which teachers integrate movement in the classroom, independent of any of these factors that may affect implementation, is because we lack empirical evidence of MI behaviors obtained through direct observation. Moreover, operational definitions of MI behaviors are needed to facilitate teacher education, teacher professional development, and future research.

Systematic observation. Systematic observation is defined by Darst, Mancini, and Zakrajsek (as quoted in van der Mars, 1989) as "a trained person following stated guidelines and procedures to observe, record, and analyze interactions with the assurance that others viewing the same sequence of events would agree with his [or her] recorded data" (p. 6). Systematic observation is a proven method of capturing contextual and behavioral variables that are useful in operationally defining, advancing, and evaluating best practices in teaching (Flanders, 1970; Flanders, 1976; van der Mars, 1989) and PA promotion in a number of settings, such as physical education (McKenzie, Sallis, & Nader, 1992), afterschool programs (Weaver, Beets, Webster, & Huberty, 2014), and preschools (Brown et al., 2006). An advantage of focusing only on events or behaviors that can be directly observed is the data are believed to be a more accurate account than self-reports. The purpose is to provide a permanent record of events or activities that occurred to be analyzed at a future time and is typically used in research and supervision (van der Mars, 1989). CTs' deliberate and incidental use of MI have not been objectively quantified through systematic observation. An instrument designed to systematically observe MI can be used to provide empirical evidence of what transpires in the academic classroom, measure implementation fidelity, and yield information needed to enhance future recommendations for preserve teacher education, inservice teacher training, and the development of classroom-based movement integration interventions. This information will extend the descriptive knowledge base needed to inform policy decisions and program evaluation in the context of school-wide efforts to promote children's daily PA. The steps in conducting systematic observation include deciding what behavior(s) to observe, defining the behavior(s), selecting or creating an appropriate instrument to measure the behavior(s), establishing observer reliability, conducting observations, and summarizing and interpreting the data. The principal recording strategies utilized in systematic observations are event, duration, or interval recording, or momentary time sampling (van der Mars, 1989b).

Event recording is typically appropriate for discrete behaviors or events that may happen repeatedly and yields data on frequency of occurrence. Duration recording is appropriate for examining a few discrete behaviors that are not likely to change often and provides data on temporal aspects of the observation. Common measurement units for duration recording are minutes and seconds. Interval recording involves reporting the presence or absence of an event or behavior during a predetermined period of time, or interval. Intervals are divided into equal lengths of time, usually ranging from 6 to 30 seconds, and alternate between observing and recording. Units of measurement for this type of recording are frequency of intervals, which is usually later converted to a percentage of total intervals. Momentary time sampling is similar to interval recording in that the observation is divided into equal intervals of time. However, unlike interval recording where the observation takes place *during* the interval, momentary time sampling requires the observation to take place *at the end* of the interval. Data collected using this strategy are reported as a percentage of total intervals. In designing a systematic observation instrument, it is useful to have a conceptual framework that guides the development of initial observation categories.

MI conceptual framework. For the purpose of this review, movement integration (MI) is defined globally as any strategy CTs utilize to increase classroom PA opportunities or decrease non-recreational sedentarism for their students (IOM, 2013). Recent recommendations for MI focus on two major strategies: (a) incorporating PA breaks between academic lessons, and (b) infusing PA into academic lessons (Webster et al., 2015). These strategies reflect ways CTs can deliberately integrate movement and are consistent with national recommendations for classroom PA as part of a whole-of-school approach to PA (AAHPERD, 2011; CDC, 2013; IOM, 2013; NASPE, 2013). However, there are also a few ways CTs integrate movement in a more subtle manner. These can be thought of as incidental opportunities. An opportunity for movement was present; however, it was not explicitly driven by the teacher. For example, a teacher may establish a procedure for students to walk around the perimeter of the classroom each time they need to sharpen a pencil. The rule or procedure was directed by the teacher when it was initially established; however, the procedure may no longer be teacher directed when it is observed because it has become an established routine. In this case, the opportunity to move (i.e., walking around the perimeter of the room) is considered incidental. Another example of incidental MI is when the classroom is arranged in a particular way to facilitate movement. Again, this is not teacher directed each time; however, when a student moves as a result of the way the furniture or fixtures were previously arranged, it is considered incidental. Whether deliberate or incidental, MI results in an opportunity for students to not be sedentary, regardless of the intensity level of the movement (Webster, et al., 2015). This means students can be engaged in light-, moderate-, or vigorous-intensity PA (IOM, 2013).

Direct observation has been used in the elementary classroom within PA research but only to measure the intensity level of student PA (Donnelly et al., 2009). The only evidence of any direct observation of MI is from one observational study, conducted in New York City schools, of a classroom-based PA program called Move-To-Improve (Dunn et al., 2010). Although trained data collectors conducted full-day observations in the elementary classroom and recorded information related to movement integration, there was no evidence presented that specific coding rules and procedures were followed, nor that a specific systematic observation instrument was developed, adapted, or employed (Dunn et al., 2010).

There is a lack of empirical evidence documenting what teachers are doing to integrate movement in the academic classroom. For example, only one instance of incidental MI is even documented in the literature (Janulewicz, 2008). In addition, measurements of intervention implementation fidelity are needed to in order to ensure compliance and that interventions are delivered as designed. Part of the reason we have limited knowledge about current MI practices is because we lack objective evidence obtained through direction observation. The limited information available has relied on self-report and results from the absence of a systematic observation tool that captures MI.

Summary

Extant literature supports the importance of PA for children, including physical and mental health benefits. It also demonstrates an increasing concern about the sedentary state of children, detrimental effects of sedentarism, and the need to decrease non-recreational sedentary opportunities. Schools have historically played a role in children's health and continue to be recommended as settings to intervene.

Within schools, the classroom setting receives support from the literature as a place to target non-recreational sedentarism and to do so through MI (Bartholomew & Jowers, 2011; Erwin et al., 2011; Holt et al., 2013; Webster et al., 2015). Movement integration offers CTs a way to incorporate PA in their classroom, thereby contributing to students' progress toward meeting daily PA recommendations (Bassett et al., 2013), without compromising academic performance (IOM, 2013). In fact, MI offers additional benefits including on-task behavior improvements (Grieco, Jowers, & Bartholomew, 2009; Howie, 2013; Mahar et al., 2006; Mahar, 2011), positive affect (Howie et al., 2014), and cognitive function (Donnelly & Lambourne, 2011; Elmakis, 2010; Howie et al., 2014). Unfortunately, despite these advantages of movement integration, there is limited evidence of implementing PA in the classroom, and that is partly due to a lack of objective measurements.

Objective measurements of MI are needed for several reasons. Objective measures of MI implementation will help advance the descriptive knowledge base that is currently quite small and limited by a reliance on self-report data. Objective measures will enable implementation fidelity to be evaluated to ensure MI is being implemented as designed and also enhance research on implementation fidelity by contributing another perspective to the current discussion centered on self-reported data. Further, monitoring implementation provides a way to identify and address any variables related to MI implementation (e.g. teacher training, physical space constraints) that may impact effectiveness.

One type of objective instrumentation that has been proven to capture contextual and behavioral factors used to advance best practices in education and teacher education is systematic observation (Flanders, 1970; Flanders, 1976). Systematic observation tools have also been used successfully in a variety of PA and PE contexts (McKenzie et al., 1992; Weaver et al., 2014) making it an attractive possibility for measuring MI. Further, the instrument development process will create operational definitions of MI behaviors providing a common language that can be used for educational and research purposes. Developing a systematic observation instrument able to capture MI will address the need for objective measurement, and at the same time, help provide empirical evidence of CSPAP effectiveness that is needed to support advancing adoption of these programs.

CHAPTER 3: STUDY 1

SYSTEMATIC REVIEW AND META-ANALYSIS OF MULTI-COMPONENT INTERVENTIONS THROUGH SCHOOLS TO INCREASE PHYSICAL ACTIVITY[1]

[1]Russ, L., Webster, C.A., Beets, M.W., & Phillips, D.S. Accepted by *Journal of Physical Activity and Health.*

Reprinted here with permission of publisher.

Half of America's youth do not meet the national guideline of 60 minutes or more of moderate or vigorous-intensity physical activity (PA) each day (United States Department of Health and Human Services [USDHHS], 2008). Schools are advocated as a key setting for helping youth to meet this recommendation (Centers for Disease Control and Prevention [CDC], 2013; Institute of Medicine [IOM], 2013; National Physical Activity Plan, 2012; Pate et al., 2006). Schools have unparalleled access to most youth for many hours on most days of the week, offer an existing infrastructure for PA promotion, have historically played an important role in promoting children's health, and can improve children's health and education through PA (IOM, 2013). Unfortunately, the effects of school-based PA interventions on the total daily PA of youth have been negligible (Metcalf, Henley, & Wilkin, 2012). These results could be due to poor delivery or uptake of intervention components, the use of insufficiently intense physical activities, or poorly timed PA sessions that merely replaced opportunities during which participants would have been equally active (Metcalf, et al., 2012).

The minimal impact of previous interventions may also be related to the quantity and quality of intervention components designed to increase PA. Recent guidelines call for a "whole-of-school" approach, which is defined as "all of a school's components and resources operat[ing] in a coordinated and dynamic manner to provide access, encouragement, and programs that enable all students to engage in vigorous- or moderate-intensity physical activity 60 minutes or more each day" (IOM, 2013, p. 367). In accordance with this approach, comprehensive school physical activity programs (CSPAP) are recommended. Distinct components of a CSPAP include (a) quality physical education (QPE), (b) PA during the school day (PADSD), (c) PA before or after school (PABAS), (d) staff wellness (SW), and (e) family and community engagement (FCE) (see Table 1) (CDC, 2013). The purpose of a CSPAP is to increase the quantity and quality of PA opportunities through schools to maximize participation in PA.

The extent to which interventions reflect, or have adopted, a whole-of-school approach remains unclear. Distilling the effects of interventions targeting multiple CSPAP components may provide a unique, and possibly more promising, perspective of extant efforts to increase youth PA through schools. The present study examined the effectiveness of multi-component interventions on increasing the total daily PA of youth. Specifically, a systematic review and meta-analysis, using the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines (Liberati et al., 2009; Moher, Liberati, Tetzlaff, & Altman, 2009) was conducted on interventions that included two or more CSPAP components.

Evidence Acquisition

Search Strategy

Studies were identified and analyzed between August 2013 and January 2014. Two reviewers conducted independent searches using two electronic databases (GoogleScholar and PubMed) and the following combinations of keywords: physical activity, school, and int*, exp*, or trial. After an initial list was generated, researchers conferred to verify the same number of hits from each database with 100% agreement. Results from each electronic database search were further analyzed by title and abstract according to the PRISMA (Liberati, et al., 2009; Moher et al., 2009) guidelines. Existing review articles on youth PA interventions were also identified and their references searched for inclusion of studies.

Inclusion Criteria

Interventions were included in this review that (1) occurred in the United States; (2) targeted any school grade level K-12; (3) were an intervention (not restricted to randomized controlled trials); (4) included two or more components reflective of the CSPAP model, with at least one targeting PA of children at their own school, during regular school hours; and (5) reported changes in total daily PA.

Assessment of Quality

A quality indicator index was developed based on previous research (Campbell, Waters, O'Meara, & Summerbell, 2001; Engbers, van Poppel, Chin A Paw, & van Mechelen, 2005; Flodmark, Marcus, & Britton, 2006; Metcalf et al., 2012; van Sluijs, McMinn, & Griffin, 2007; van Sluijs, van Poppel, & van Mechelen, 2004), adaptations from the Cochrane tool (Higgins, Green, & Collaboration, 2008), and researcher input to describe elements of quality for each study included in this review (Table 2). Two researchers conducted independent evaluations of study quality. Disagreements were resolved through discussion. When needed, a third researcher was enlisted to reach consensus. If definitive evidence of an indicator was absent, the ground rule established was to report it as unknown in order to avoid making assumptions.

Data Extraction

Three reviewers extracted from each study information regarding study design, participants, sample size, length of intervention, descriptions of each component aligned with the CSPAP model, PA measurements, outcome scores, context, and risk of bias. Risk of bias was identified as blinding of participants/personnel (performance bias) and blinding of outcome assessment (detection bias) based on items from the Cochrane Group's tool (Higgins et al., 2008). When extracting information about QPE, the intervention had to present a minimum of one element of QPE (see Table 1). A decision was made to separate staff implementation and staff wellness as two subcomponents within the staff involvement component of the CSPAP model. This was because all of the interventions included staff involvement (i.e., staff were trained/encouraged to promote PA or were a part of implementation). The inclusion of staff implementation would have altered the range of components across interventions and would not have helped discriminate program effectiveness by CSPAP components. Descriptive information of each subcomponent was extracted for qualitative purposes. Three reviewers conferred on each study to determine the relevant outcome measures for the meta-analysis. One reviewer extracted data on total PA, MVPA, vigorous physical activity (VPA), and sedentary activity levels (e.g., minutes of MVPA, MET-weighted minutes of MVPA, step counts, number of 30-minute blocks/day, accelerometer counts, change scores, adjusted odds ratios, and energy expenditures). Data were extracted in the units reported. Most often, means, standard deviations, and p-values were presented and extracted. When change scores were present, the difference and p-values were extracted.

Calculating Effectiveness (Meta-Analysis)

Standardized mean difference effect sizes were calculated on all PA behaviors reported on a daily basis for each study. Examples of these include minutes of PA per day, number of 30 minute blocks of MVPA per day, or number of past 7 days with 60 or more minutes of PA. Each study's daily PA effects were extracted and transformed into a common metric (i.e., Hedge's g) based on the study design, as well as, the amount of information provided in the published article (Morris & DeShon, 2002). Where insufficient numerical data were provided in the published article, requests were made to the study's primary author to obtain the necessary information. Only a single article was excluded where insufficient information was presented in the published study and attempts to obtain additional information were unsuccessful (Story et al., 2012).

Various PA measurements, such as self-report instruments, objective monitors (e.g., pedometers or accelerometers), and different protocols to distill information from the same measure (e.g., accelerometer minutes of MVPA per day versus average daily counts per minute), were used to quantify the intervention-related effects for changes in daily PA across the included studies. Because of this, the assumption was made that each of these instruments was measuring an aspect of the construct of daily PA, and therefore, were pooled together in the analyses.

For each study, individual effect sizes and corresponding 95% CIs were calculated for each outcome measure. A single study reported effects from a 3-arm intervention that included a control condition and two intervention conditions (Sallis et al., 1997). This single study was considered as two separate studies (comparisons between control and intervention one and two, separately) in all subsequent analyses. Hedge's g (Hedges, 1982) was used to adjust effect size estimates for small sample sizes by multiplying the effect size with the correction factor (1–(3/[4N–9])) (where N is the total sample size at the child level). For the analytical models, all pooled effects weighted the contribution of each study by the study's standard deviation and sample size and used the study as the unit of analysis. Pooled effect sizes were calculated using a random-effects inverse variance (DerSimonian & Laird, 1986) model based on the assumption that all studies were estimating different, yet related, treatment effects (i.e., all studies were intervening on youth daily PA). The percentage of the total variability in an effect size due to heterogeneity (between-studies variability) was estimated with I-squared (I^2) (Huedo-Medina, Sanchez-Meca, Marin-Martinez, & Botella, 2006). The percentages associated with I^2 are interpreted as low (25%), medium (50%), and high (75%) heterogeneity (i.e., between study variability), respectively.

A series of models were estimated based on the following. First, an overall pooled Hedge's g was estimated across all studies to determine the overall effect of the interventions on youth daily PA. Second, pooled effects were compared across the interventions by the individual CSPAP components described in the interventions (e.g., studies that included QPE vs. not including QPE), as well as, the total number of CSPAP components reflected in the intervention (range 2 to 4). These models were also compared across potential moderators of intervention effectiveness. The moderators were objective versus non-objective measures of daily PA and gender (reporting boys and/or girls, separately, or reporting boys and girls combined). Two studies (Caballero et al., 2003; Sallis et al., 1997) included both objective and self-report measures of daily PA and were treated separately in the analysis comparing effect sizes between measurement types. For the evaluation of gender on the effect size, the effect was pooled at the gender level for studies that reported two or more measured PA outcomes for boys and/or girls. For instance, a study (e.g., Sallis et al., 1997) could report changes in daily PA via accelerometry and also include changes in PA captured via self-report for boys and girls, separately. Sensitivity analyses were conducted on the pooled estimates to determine the influence of any given study's results on the overall effect size by omitting one study and re-estimating the pooled effect sizes. Finally, meta-regression was used to evaluate the impact of study length on the estimated effect sizes. All analyses were conducted using Comprehensive Meta Analysis (v.2.2.048).

Evidence Synthesis

Literature Search

A total of 1,087 records were identified and the abstracts screened by three reviewers. Of these, 359 full-text documents were identified for inclusion. Disagreements were discussed until consensus was reached. A final count of 14 unique studies (Caballero et al., 2003; Gortmaker et al., 1999; Hoelscher et al., 2010; Luepker et al., 1996; Neumark-Sztainer, 2003; Neumark-Sztainer et al., 2010; Pate et al., 2005; Sallis, et al., 1997; Sallis et al., 2003; Seo et al., 2013; Springer et al., 2012; Webber et al., 2008; Williamson et al., 2007; Young, Phillips, Yu, & Haythornthwaite, 2006) met inclusion criteria and were included in this review (Figure 1).

Descriptions of CSPAP Components

Across the 14 studies, a total of 51,560 participants from 307 schools ranging in mean age from 7.0 (Caballero et al., 2003) -15.8 (Neumark-Sztainer, et al., 2010) years old were included in baseline data collection (Table 3). Eleven (Caballero et al., 2003; Gortmaker et al., 1999; Luepker et al., 1996; Neumark-Sztainer et al., 2003; Neumark-Sztainer et al., 2010; Pate et al., 2005; Sallis et al., 1997; Sallis et al., 2003; Webber et al., 2008; Williamson et al., 2007; Young et al., 2006) out of 14 were randomized controlled studies, with the average number of schools and students across all studies as 21 (range 1 (Young et al., 2006) to 96 (Luepker et al., 1996)) and 3,683 (range 201(Neumark-Sztainer et al., 2003) to 26,616,(Sallis et al., 2003)), respectively. The median intervention length and sample size was 360 days (Gortmaker et al., 1999; Neumark-Sztainer et al., 2010; Sallis et al., 1997; Sallis et al., 2003; Webber et al., 2008; Williamson et al., 2007) and 1099 students, respectively. Five of the interventions (Neumark-Sztainer et al., 2003; Neumark-Sztainer, et al. 2010; Pate et al., 2005; Webber et al., 2008; Young et al., 2006) focused solely on females, 13 (Caballero et al., 2003; Gortmaker et al., 1999; Hoelscher et al., 2010; Luepker et al., 1996; Neumark-Sztainer, 2003; Neumark-Sztainer et al., 2010; Pate et al., 2005; Sallis, et al., 1997; Sallis et al., 2003; Seo et al., 2013; Springer et al., 2012; Williamson et al., 2007; Young et al., 2006) employed self-report as a measure of physical activity, one study(Webber et al., 2008) did not use self-reported measures (direct observation and motion sensor), and 2 studies (Caballero et al., 2003; Sallis et al., 1997) used a combination of both self-report and objective measurement. No study included all five CSPAP components. The median number of CSPAP components was 2 with a range of 2 (Luepker et al., 1996; Neumark-

Sztainer et al., 2003; Neumark-Sztainer et al., 2010; Pate et al., 2005; Sallis et al., 1997) - 4 (Sallis et al., 2003) intervention components. The most common components observed were FCE (n=14) (Caballero et al., 2003; Gortmaker et al., 1999; Hoelscher et al., 2010; Luepker et al., 1996; Neumark-Sztainer et al., 2003; Neumark-Sztainer et al., 2010; Pate et al., 2005; Sallis et al., 1997; Sallis et al., 2003; Seo 2013; Springer et al., 2012; Webber et al., 2008; Williamson et al., 2007; Young et al., 2006) followed by QPE (n=12) (Caballero et al., 2003; Gortmaker et al., 1999; Hoelscher et al., 2010; Luepker et al., 1996; Neumark-Sztainer et al., 2003; Neumark-Sztainer et al., 2010; Pate et al., 2005; Sallis et al., 1997; Sallis et al., 2003; Seo et al., 2013; Webber et al., 2008; Young et al., 2006). PA During the School Day (Caballero et al., 2003; Hoelscher et al., 2010; Sallis et al., 2003; Springer et al., 2012; Williamson et al., 2007) was included in less than half of the studies (n=5). Components represented the least were SW (Gortmaker et al., 1999; Seo et al., 2013) (n=2) and PABAS (Sallis et al., 2003) (n=1).

Quality physical education. Twelve studies (Caballero et al., 2003; Gortmaker et al., 1999; Hoelscher et al., 2010; Luepker et al., 1996; Neumark-Sztainer et al., 2003; Neumark-Sztainer et al., 2010; Pate et al., 2005; Sallis et al., 1997; Sallis et al., 2003; Seo et al., 2013; Webber et al., 2008; Young et al., 2006) incorporated components reflecting QPE as previously defined (Table 1). The most common approach was to increase PA in physical education (PE) (Caballero et al., 2003; Gortmaker et al., 1999; Hoelscher et al., 2010; Luepker et al., 1996; Neumark-Sztainer et al., 2003; Neumark-Sztainer et al., 2010; Pate et al., 2005; Sallis et al., 1997; Sallis et al., 2003; Seo et al., 2013; Webber et al., 2008; Young et al., 2006) whether by increasing energy expenditure or by replacing inactive time with active time. Other common pieces of QPE evident in the interventions included increasing enjoyment (Neumark-Sztainer et al., 2003; Neumark-Sztainer et al., 2010; Pate et al., 2005; Sallis et al., 1997), self-efficacy(Neumark-Sztainer et al., 2003; Neumark-Sztainer et al., 2010; Pate et al., 2005), providing equipment/supporting purchase of for PE programs(Webber et al., 2008), and developing movement skills (Pate et al., 2005; Sallis et al., 1997; Webber et al., 2008; Young et al., 2006).

PA during the school day. Five studies (Caballero et al., 2003; Hoelscher et al., 2010; Sallis et al., 2003; Springer et al., 2012; Williamson et al., 2007) included this component. PADSD occurred in the academic classroom (Caballero et al., 2003; Hoelscher et al., 2010; Williamson et al., 2007), at recess (Caballero et al., 2003; Sallis et al., 2003; Springer et al., 2012; Williamson et al., 2007), and during drop-in sessions where equipment was provided after lunch and students could choose to be active (Sallis et al., 2003). One of the most common approaches to including PA during the school day was in the academic classroom where 3 (Caballero et al., 2003; Hoelscher et al., 2010; Williamson et al., 2007) out of the 5 studies included PA breaks in the classroom. Two studies (Sallis et al., 2003; Williamson et al., 2007) promoted PA during the school day by providing equipment for students to use. While PA during the school day occurred in a variety of settings, the majority of interventions (Caballero et al., 2003; Hoelscher et al., 2010; Williamson et al., 2007) with this component specifically required the help of the classroom teacher.

PA before and after school. Only one study (Sallis et al., 2003) included a before or after school physical activity component to the intervention. The strategy included environmental and policy changes to impact student PA levels, such as policy changes to allow students access to facilities for PA after school and hiring personnel to facilitate PA programs.

Staff wellness. Two interventions (Gortmaker et al., 1999; Seo et al., 2013) included a staff wellness component; however, details about the specific opportunities were limited. Gortmaker et al.(1999) took teacher/staff interests into consideration and offered wellness sessions delivered by outside agencies, but Seo et al.(2013) did not provide any details of the wellness events.

Family and community engagement. All fourteen studies (Caballero et al., 2003; Gortmaker et al., 1999; Hoelscher et al., 2010; Luepker et al., 1996; Neumark-Sztainer et al., 2003; Neumark-Sztainer et al., 2010; Pate et al., 2005; Sallis et al., 1997; Sallis et al., 2003; Seo et al., 2013; Springer et al., 2012; Webber et al., 2008; Williamson et al., 2007; Young et al., 2006) included a component designed to engage families and/or communities in promoting youth PA. For example, a common way to engage families was to increase communication with families (Pate et al., 2005; Webber et al., 2008). This took the form of sending home newsletters with information about healthy lifestyle habits(Caballero et al., 2003; Gortmaker et al., 1999; Neumark-Sztainer et al., 2003; Neumark-Sztainer et al., 2010; Sallis et al., 1997; Sallis et al., 2003; Seo et al., 2013; Williamson et al., 2007; Young et al., 2006), sending home specific adult-child homework assignments (Gortmaker et al., 1999; Luepker et al., 1996; Sallis et al., 1997; Young et al., 2006), and offering formal parental education approaches (Sallis et al., 2003; Williamson et al., 2007; Young et al., 2006) including Internet-based education programs (Williamson et al., 2007) and workshops (Young, et al., 2006). A variety of strategies for family and community involvement were represented ranging from educational events (Williamson et al., 2007) to health fairs (Seo et al., 2013) and active events (Hoelscher et al., 2010; Luepker et al., 1996; Pate et al., 2005), such as family fun nights. Community involvement was often described as partnering with schools in conducting and promoting events (e.g., Family Fun Nights and promoting recreation center activity programs) (Hoelscher et al., 2010; Sallis et al., 2003); Springer et al., 2012; Webber et al., 2008), with a few studies (Neumark-Sztainer et al., 2003; Neumark-Sztainer et al., 2010) helping students connect to PA opportunities in the community. For example, New Moves (Neumark-Sztainer et al., 2003; Neumark-Sztainer et al., 2010) was an intervention specifically designed to link the all-girls PE class with opportunities for PA outside of school. Community guests were invited into PE class to lead new and unique physical activities for the girls to try (e.g., kickboxing, yoga or water aerobics), and students took field trips to community centers where they could see how, where, and in what ways, they could be physically active outside of school.

Meta-Analysis

A total of 40 effects were extracted from the 14 studies (Caballero et al., 2003; Gortmaker et al., 1999; Hoelscher et al., 2010; Luepker et al., 1996; Neumark-Sztainer et al., 2003; Neumark-Sztainer et al., 2010; Pate et al., 2005; Sallis et al., 1997; Sallis et al., 2003; Seo et al., 2013; Springer et al., 2012; Webber et al., 2008; Williamson et al., 2007; Young et al., 2006) that described 15 interventions and were used in the analytical models. The results of the analyses are presented in Table 4 and in Figure 2. The overall effect of the interventions on youth total daily PA was minimal, with a pooled effect size of $g = 0.11$ (95% Confidence Interval [95CI] 0.03 to 0.19). Comparable effects were observed for studies that reported daily physical activity for boys ($g = 0.09$, 95CI -0.10 to

0.28) and girls (g = 0.11, 95CI -0.02 to 0.23), separately, and in studies that reported boys and girls combined (g = 0.12, 95CI 0.05 to 0.19). Across all studies and by studies reporting gender specific activity outcomes, as the number of CSPAP components included in the intervention increased, the effect size associated with the change in daily physical activity increased from 0.06 to 0.19 to 0.29 for 2 (Luepker et al., 1996; Neumark-Sztainer et al., 2003; Neumark-Sztainer et al., 2010; Pate et al., 2005; Sallis et al., 1997; Springer et al., 2012; Webber et al., 2008; Williamson et al., 2007; Young et al., 2006), 3,(Caballero et al., 2003; Gortmaker et al., 1999; Hoelscher et al., 2010; Seo et al., 2013) and 4 (Sallis et al., 2003) components present, respectively – however, only a single study (Sallis et al., 2003) included 4 CSPAP components. Studies that employed objective measures of physical activity (Caballero et al., 2003; Sallis et al., 1997; Webber et al., 2008) exhibited smaller effect sizes than studies using self-report measures of physical activity (Caballero et al., 2003; Gortmaker et al., 1999; Hoelscher et al., 2010; Luepker et al., 1996; Neumark-Sztainer et al., 2003; Neumark-Sztainer et al., 2010; Pate et al., 2005; Sallis et al., 1997; Sallis et al., 2003; Seo 2013; Springer et al., 2012; Williamson et al., 2007; Young et al., 2006) (0.02 vs. 0.12). Evaluation of the inclusion of each specific CSPAP component found that studies that included PADSD (0.19 vs. 0.07), PABAS (0.29 vs. 0.10), and SW (0.21 vs. 0.09) were associated with larger effect sizes than studies that did not include these components. The only CSPAP component associated with a smaller effect size was QPE (0.10 vs. 0.16). Results from the meta-regression found that study length had no effect on overall study effect size. Based on the sensitivity analyses, two studies (Sallis et al., 2003; Seo et al., 2013) upwardly influenced the overall pooled effects across all studies (0.07 vs. 0.11), and for studies that reported boys (0.01 vs. 0.09) and girls (0.02 vs. 0.11), separately. A single study (Luepker, et al., 1996) downwardly influenced the overall effect size for studies reporting boys and girls combined (0.12 vs. 0.15).

Discussion

Main Findings

While schools are recommended as a key setting for increasing youth PA, school-based interventions have been minimally effective (Metcalf et al., 2012). However, previous reviews have not distilled the effectiveness of multi-component interventions, in light of recommendations calling for a whole-of-school approach (i.e., CSPAPs) to PA promotion (CDC, 2013; IOM, 2013). This review uniquely considered the effectiveness of interventions including two or more CSPAP components. Unfortunately, the results suggest multi-component interventions have had minimal impact on the total daily PA of youth.

The reasons for the lack of effectiveness are unclear. The intervention components offering the most insight into effectiveness are QPE, PADSD, and staff involvement. Counter to expectations, the interventions with QPE components were associated with a smaller effect size than ones without QPE (0.10, 0.16). Despite evidence showing positive changes in PA within PE when targeted in isolation (Lonsdale et al., 2013), the types of strategies used in the interventions herein to increase youth PA in PE have, as a whole, not added to the capacity of multi-component approaches for increasing total daily PA. In the included interventions, PE was delivered by a variety of people including project staff, classroom teachers, and certified specialists, and did not meet guidelines for weekly allocated time or opportunities to learn, contradicting current CSPAP recommendations for QPE (CDC, 2013). Specifically, PE should be delivered by a qualified PE teacher and classes should meet for at least 150 minutes per week (elementary school) or 225 minutes per week (middle and secondary school) (IOM, 2013). Ensuring the characteristics of QPE are incorporated when including this component in multi-component approaches should help to maximize intervention effectiveness.

While interventions that included PADSD contributed to a greater effect size than interventions without this component (0.19 vs. 0.10), there may be untapped potential to maximize its effectiveness. This is consistent with findings that the contexts for PADSD, such as recess (Ridgers, Stratton, & Fairclough, 2006) and the academic classroom (Bartholomew & Jowers, 2011; Mahar et al., 2006) can be used for increasing PA (IOM, 2013). While these contexts are promising avenues, only three studies (Caballero et al., 2003; Hoelscher et al., 2010; Williamson et al., 2007) employed classroom activity breaks. Providing equipment for students to use (Sallis et al., 2003; Young et al., 2006) and scheduling PA time during and after lunch (Sallis et al., 2003; Springer et al., 2012) are important strategies, but used in isolation are not enough to maximize the potential effectiveness of this component. A more coordinated approach employing more than one of these strategies across contexts may increase effectiveness of this component.

Examining staff involvement (specifically, the subcomponent staff implementation) revealed inconsistencies in staff training. The information reported showed staff training was more often a one-shot professional development session instead of ongoing professional learning opportunities. Notable exceptions included

Sallis et al. (2003) which included five 3-hour training sessions for PE teachers and Sallis et al. (1997) which included over 32 hours of training for classroom teachers across 7 sessions. Few training sessions (Sallis et al., 2003; Williamson et al., 2007) demonstrated collaboration across school components and resources, contradicting current recommendations for a whole-of-school approach (IOM, 2013). This component can maximize the effectiveness of multi-component interventions by using ongoing professional learning opportunities that occur within the school community, incorporating experiential learning, collaborating with other areas of the school, and providing resources and equipment with which teachers are familiar(Till & Ferkins, 2014).

Overall, the interventions provided little information across components, which limited understanding of exactly what took place. For example, information about the supervision of PADSD opportunities, the physical space provided for PA, or the number of opportunities actually presented to students to be active was missing. We can only speculate about the extent to which recess was supervised, how frequently adults encouraged students to be active (if at all), and what such promotion behaviors looked like, even though these elements align with current recommendations for maximizing PADSD (CDC, 2013). Additionally, we know very little about the actual implementation of activity breaks in the classroom and structured time throughout the day. While teachers may have received training on how to incorporate these elements throughout their day to promote PA, there is little evidence to confirm PADSD implementation occurred as designed. Knowing how classroom teachers are promoting PA to students can also provide information useful for future intervention design. For example, giving teachers the flexibility to individualize strategies for PADSD to fit their environment (Till

& Ferkins, 2014), such as a focus on integrating, rather than adding, PA into already existing classroom content lessons, may strengthen the effectiveness of this component.

In many cases, details about the PE curriculum were missing. Information about the alignment with national standards, content progression, developmentally appropriate activities, and the qualifications of curriculum designers may provide insight into the limited effectiveness of this component since these are identified as components of QPE (CDC, 2013; Erwin, Beighle, Carson, & Castelli, 2013). Including certified PE teachers in the planning process for such interventions, and supporting them through ongoing professional development, may be a strategy to strengthen the elements within this component (Erwin et al., 2013; IOM, 2013).

Given that the SW component was associated with a larger effect size (0.21 vs. 0.09) than interventions without it, it would be useful to know more about the design and implementation of activities within this component. For example, information regarding the frequency, duration, and nature of SW opportunities, in addition to staff attendance/participation records, would be useful in better evaluating the effectiveness of SW strategies.

Limitations

This review searched only published studies and review articles for inclusion which may have excluded some scholarly work from initial consideration. Because total daily PA was the most common way PA was reported, only studies that reported total daily PA and had two or more components reflecting the CSPAP model were included in this review. This resulted in exclusion of interventions (e.g. PAAC (Donnelly et al., 2009), PLAY (Pangrazi, Beighle, Vehige, & Vack, 2003)) from this review but should not devalue their contribution to the field.

Recommendations

The effect of a true five-component CSPAP intervention is unknown. The increased effect size associated with the increased number of components suggests we should continue to pursue a whole-of-school approach as a potentially effective means to meaningfully increase the total daily PA of youth. However, we may not yet know how best to maximize each component or how to harness dynamic interactions between components. Securing experts in the field to create, implement, and evaluate interventions and materials is important going forward. We recommend multi-disciplinary teams consisting of research scholars and community partners with related backgrounds to coordinate strong, community-based collaborative approaches.

Results of the sensitivity analysis indicated two studies (Sallis et al., 2003; Seo et al., 2013) contributed significantly to the overall effect size. Both studies used policy development or change to facilitate increased opportunities for student PA. Sallis et al. (2003) involved staff through a series of staff development sessions (five 3-hour sessions) focused on teacher instructional skills and implementing new curricula to increase student PA. It was also the only study in this review that included four of the five components of the CSPAP model and one of the only studies (Sallis et al., 2003; Williamson et al., 2007) to demonstrate collaboration across components of a school which is consistent with the whole-of-school approach.

Conclusion

Results suggest that taking a multi-component approach to increasing youth PA is an appropriate path, but strategies within and across components may need to be reconsidered for maximal impact. Current guidelines describing a whole-of-school approach and CSPAPs offer relevant frameworks that merit investigation. Future interventions that reflect all five components of the CSPAP model, align with current recommendations, provide detailed descriptions of intervention component design and implementation, and demonstrate dynamic collaboration across all five components are needed.

References

Bartholomew, J. B., & Jowers, E. M. (2011). Physically active academic lessons in elementary children. *Preventive Medicine, 52*, S51-S54.

Caballero, B., Clay, T., Davis, S. M., Ethelbah, B., Rock, B. H., Lohman, T., . . . & Stephenson, L. (2003). Pathways: a school-based, randomized controlled trial for the prevention of obesity in American Indian school children. *The American Journal of Clinical Nutrition, 78*(5), 1030-1038.

Campbell, K., Waters, E., O'Meara, S., & Summerbell, C. (2001). Interventions for preventing obesity in childhood. A systematic review. *Obesity Reviews, 2*(3), 149-157.

Centers for Disease Control and Prevention [CDC]. (2013). Comprehensive School Physical Activity Programs: A guide for schools. Retrieved from http://www.cdc.gov/healthyyouth/physicalactivity/cspap.htm.

DerSimonian, R., & Laird, N. M. (1986). Meta-analysis in Clinical Trials. *Controlled Clinical Trials, 7*, 177-178.

Donnelly, J. E., Greene, J. L., Gibson, C. A., Smith, B. K., Washburn, R. A., Sullivan, D. K., . . . & Ryan, J. J. (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.

Engbers, L. H., van Poppel, M. N., Chin A Paw, M. J., & van Mechelen, W. (2005). Worksite health promotion programs with environmental changes: a systematic review. *American Journal of Preventive Medicine, 29*(1), 61-70.

Erwin, H., Beighle, A., Carson, R. L., & Castelli, D. M. (2013). Comprehensive School-Based Physical Activity Promotion: a review. *Quest, 65*(4), 412-428.

Flodmark, C., Marcus, C., & Britton, M. (2006). Interventions to prevent obesity in children and adolescents: a systematic literature review. *International Journal of Obesity, 30*(4), 579-589.

Gortmaker, S. L., Peterson, K., Wiecha, J., Sobol, A. M., Dixit, S., Fox, M. K., & Laird, N. (1999). Reducing obesity via a school-based interdisciplinary intervention among youth: Planet Health. *Archives of Pediatrics & Adolescent Medicine, 153*(4), 409.

Hedges, L. V. (1982). Estimation of effect size from a series of independent experiments. *Psychological Bulletin, 92*, 490-499.

Higgins, J. P., Green, S., & Collaboration, C. (2008). *Cochrane Handbook for Systematic Reviews of Interventions* (Vol. 5): Wiley Online Library.

Hoelscher, D. M., Springer, A. E., Ranjit, N., Perry, C. L., Evans, A. E., Stigler, M., & Kelder, S. H. (2010). Reductions in child obesity among disadvantaged school children with community involvement: the Travis County CATCH Trial. *Obesity, 18,* S36-S44.

Huedo-Medina, T. B., Sanchez-Meca, J., Marin-Martinez, F., & Botella, J. (2006). Assessing heterogeneity in meta-analysis: Q statistic or I2 index? *Psychological Methods, 11*(2), 193-206.

Institute of Medicine [IOM]. (2013). *Educating the student body: Taking physical activity and physical education to school*. Washington, DC: The National Academies Press.

Liberati, A., Altman, D. G., Tetzlaff, J., Mulrow, C., Gøtzsche, P. C., Ioannidis, J. P., . . . & Moher, D. (2009). The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *Annals of Internal Medicine, 151*(4), W-65-W-94.

Lonsdale, C., Rosenkranz, R. R., Peralta, L. R., Bennie, A., Fahey, P., & Lubans, D. R. (2013). A systematic review and meta-analysis of interventions designed to increase moderate-to-vigorous physical activity in school physical education lessons. *Preventive Medicine, 56*(2), 152-161.

Luepker, R. V., Perry, C. L., McKinlay, S. M., Nader, P. R., Parcel, G. S., Stone, E. J., . . . & Johnson, C. C. (1996). Outcomes of a field trial to improve children's dietary patterns and physical activity. *Journal of the American Medical Association, 275*(10), 768-776.

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.

Metcalf, B., Henley, W., & Wilkin, T. (2012). Effectiveness of intervention on physical activity of children: systematic review and meta-analysis of controlled trials with objectively measured outcomes (EarlyBird 54). *British Medical Journal, 345*.

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Annals of Internal Medicine, 151*(4), 264-269.

Morris, S. B., & DeShon, R. P. (2002). Combining effect size estimates in meta-analysis with repeated measures and independent-groups designs. *Psychological Methods, 7*(1), 105-125.

National Physical Activity Plan Alliance. (2012). National Physical Activity Plan.

Neumark-Sztainer, D., Story, M., Hannan, P. J., & Rex, J. (2003). New Moves: a school-based obesity prevention program for adolescent girls. *Preventive Medicine, 37*(1), 41-51.

Neumark-Sztainer, D. R., Friend, S. E., Flattum, C. F., Hannan, P. J., Story, M. T., Bauer, K. W., . . . & Petrich, C. A. (2010). New Moves—preventing weight-related problems in adolescent girls: a group-randomized study. *American Journal of Preventive Medicine, 39*(5), 421-432.

Pangrazi, R. P., Beighle, A., Vehige, T., & Vack, C. (2003). Impact of Promoting Lifestyle Activity for Youth (PLAY) on children's physical activity. *Journal of School Health, 73*(8), 317-321.

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

Pate, R. R., Ward, D. S., Saunders, R. P., Felton, G., Dishman, R. K., & Dowda, M. (2005). Promotion of physical activity among high-school girls: a randomized controlled trial. *Journal Information, 95*(9).

Ridgers, N. D., Stratton, G., & Fairclough, S. J. (2006). Physical activity levels of children during school playtime. *Sports Medicine, 36*(4), 359-371.

Sallis, J. F., McKenzie, T. L., Alcaraz, J. E., Kolody, B., Faucette, N., & Hovell, M. F. (1997). The effects of a 2-year physical education program (SPARK) on physical activity and fitness in elementary school students. Sports, Play and Active Recreation for Kids. *American Journal of Public Health, 87*(8), 1328-1334.

Sallis, J. F., McKenzie, T. L., Conway, T. L., Elder, J. P., Prochaska, J. J., Brown, M., . . . & Alcaraz, J. E. (2003). Environmental interventions for eating and physical activity: a randomized controlled trial in middle schools. *American Journal of Preventive Medicine, 24*(3), 209-217.

Seo, D.-C., King, M., Kim, N., Sovinski, D., Meade, R., Lederer, A. (2013). Predictors for moderate- and vigorous-intensity physical activity during an 18-month coordinated school health intervention. *Preventive Medicine*. doi: http://dx.doi.org/10.1016.j.ypmed.2013.06.024

Springer, A. E., Kelder, S. H., Ranjit, N., Hochberg-Garrett, H., Crow, S., & Delk, J. (2012). Promoting physical activity and fruit and vegetable consumption through a community-school partnership: the effects of Marathon Kids® on low-income elementary school children in Texas. *Journal of Physical Activity & Health, 9*(5), 739-753.

Story, M., Hannan, P. J., Fulkerson, J. A., Rock, B. H., Smyth, M., Arcan, C., & Himes, J. H. (2012). Bright Start: Description and Main Outcomes From a Group-Randomized Obesity Prevention Trial in American Indian Children. *Obesity, 20*(11), 2241-2249.

Till, J., & Ferkins, L. (2014). First-hand participation: illuminating teachers' self-perceptions of physical activity delivery. *Asia-Pacific Journal of Health, Sport and Physical Education, 5*(1), 39-54.

United States Department of Health and Human Services [USDHHS]. (2008). *Physical Activity Guidelines for Americans.* Retrieved from http://www.health.gov/paguidelines

van Sluijs, E. M., van Poppel, M. N., & van Mechelen, W. (2004). Stage-based lifestyle interventions in primary care: are they effective? *American Journal of Preventive Medicine, 26*(4), 330-343.

van Sluijs, E., McMinn, A., & Griffin, S. (2007). Effectiveness of interventions to promote physical activity in children and adolescents: systematic review of controlled trials. *British Medical Journal, 335*(7622), 677-678.

Webber, L. S., Catellier, D. J., Lytle, L. A., Murray, D. M., Pratt, C. A., Young, D. R., . . . & Jobe, J. B. (2008). Promoting physical activity in middle school girls: Trial of Activity for Adolescent Girls. *American Journal of Preventive Medicine, 34*(3), 173-184.

Williamson, D. A., Copeland, A. L., Anton, S. D., Champagne, C., Han, H., Lewis, L., . . . Stewart, T. (2007). Wise Mind Project: A School-based Environmental Approach for Preventing Weight Gain in Children. *Obesity, 15*(4), 906-917.

Young, D. R., Phillips, J. A., Yu, T., & Haythornthwaite, J. A. (2006). Effects of a life skills intervention for increasing physical activity in adolescent girls. *Archives of Pediatrics & Adolescent Medicine, 160*(12), 1255.

Table 3.1 Descriptions of CSPAP components (CDC, 2013)

CSPAP Component	Description	Example
QPE	Opportunity to learn	Instruction for 150 minutes (elementary) or 225 minutes (secondary) per week
	Appropriate instruction	Delivered by certified physical education teacher
	Meaningful content	Developmentally appropriate activities and equipment
	Student and program assessment	Content reflects national standards
		Students engaged in MVPS for at least 50% of class
		Formative and summative student assessment aligned with national standards
PADSD	Encouraging students to be active	Creating active lessons
	Providing space, time, and equipment for students	Integrating PA into academic lessons
	PA breaks during/between classes	Recess before/after lunch
		Drop-in sessions in the gym before school or before/during/after lunch
PABAS	Opportunities for students to be active before and/or after the regular school day	Traditional before and after-school programs
	Making school facilities open and available to students outside regular school day	Extracurricular activities like intramural, interscholastic, or youth sports, PA clubs
		Actively travelling to and from school (walking/biking to and from school)
SW	School employee wellness opportunities	Taking responsibility for one's health and being role model for students by being physically active themselves
		Participating in PA before, during, or after school
FCE	Engaging families and communities to be active together with students	Engaging families through adult-child homework assignments, attending educational presentations and workshops, and participating in active events (e.g. Family Fun/Fit nights)
		Establishing community-based partnerships to link school PA to opportunities in the community
		Help students identify ways to be active outside of school (e.g. 5K road races or dance classes at a community recreation center)

Table 3.2 Indicators of study quality

Study	RAND	CON	LEN[a]	CG	N	DR[c]	PAM	FOL[b]	ITT	CC	BPART[c]	BPERS[c]	BOA[c]
Caballero (2003)	•	•	3 years	•	1704	295	SR MS	Ø	ITT	•	Ø	o	•
Gortmaker (1999)	•	•	2 years	•	1560	265	SR MS	Ø	ITT	•	Ø	o	Ø
Hoelscher (2010)	o	o	4 years[d]	•	1107	Ø	SR DO	Ø	Ø	•	Ø	o	Ø
Luepker (1996)	•	•	3 years	•	5106	Ø	SR DO	Ø	ITT	•	Ø	o	Ø
Neumark-Sztainer (2003)	•	•	16 weeks	o	201	11	SR	8 months	AT	•	o	o	Ø
Neumark-Sztainer (2010)	•	•	2 years	o	356	20	SR	9 months	Ø	•	Ø	o	Ø
Pate (2005)	•	•	1 year	•	2744	633	SR	Ø	ITT	•	Ø	o	Ø
Sallis (1997)	•	•	2 years	•	955	593	SR DO MS	Ø	Ø	•	Ø	o	Ø
Sallis (2003)	•	•	2 years	•	24 schools with mean enrollments of 1109 n= approximately 26,616	Ø	SR DO	Ø	Ø	•	Ø	o	o
Seo (2013)	o	o	18 months	o	1091	39.10 %	SR	Ø	AT	•	Ø	Ø	o
Springer (2012)	o	o	6 months	Ø	511	Ø	SR	Ø	Ø	•	Ø	o	Ø
Webber (2008)	•	•	2 years	Ø	2003: 1721 2005: 3504 2006: 3502	2003: 6.9% 2005: 12% 2006: 12%	DO MS	Ø	ITT	•	Ø	o	•
Williamson (2007)	•	•	2 years	Ø	661	75	SR	Ø	Ø	Ø	Ø	o	Ø
Young (2006)	•	•	8 months	•	221	11	SR DO	Ø	Ø	•	Ø	o	Ø

Note: [a] as reported; [b] identified as unknown in absence of definitive third data point; [c] adapted from Cochrane Tool(Higgins, et al., 2008); [d] first year results reported; • yes; o no; Ø unknown; AT, as treated; BPART, blinding of participants; BPERS, blinding of personnel; BOA, blinding of outcome assessment; CC, control for confounders; CG, comparable groups; CON, control condition; DO, direct observation; DR, dropout rate; FOL, follow-up; ITT, intent to treat; LEN, length of study; MS, motion sensor; N, study size; PAM, PA measure; RAND, randomization; SR, self-report

Table 3.3 Characteristics of multi-component school-based PA interventions (named reference is primary reference)

				CSPAP components							Outcomes		
							SINV						
Study	Study design	Population	Intervention	QPE	PADSD	PABAS	SI	SW	FCE	Exp	Primary outcome	PA outcome	PA instrument
Caballero (2003)	randomized controlled trial	n=1704; 41 schools; mean age 7.0-8.2 years; both sexes	Pathways	•	•		•		•	3 y	percentage body fat	total daily PA	accelerometers PAQ
Gortmaker (1999)	randomized controlled field trial	n=1560; 10 schools; mean age 11.7 (SD 0.7) years; both sexes	Planet Health	•			•	•	•	2 y	presence of obesity	total daily MVPA; hours of sedentary behavior	YAQ Food and Activity Survey Television and Video Measure
Hoelsher (2010)	serial cross-sectional	n=1107; 30 schools; mean age 9.92 (SD 0.51) years; both sexes	Travis County CATCH Project	•	•		•		•	3 y[b]	reduce presence of overweight and obesity	daily PA patterns; MVPA during PE	SOFIT SPAN questionnaire
Luepker (1996)	randomized controlled field trial	n=5106; 96 schools; mean age 8.76 years; both sexes	CATCH	•			•		•	3 y	increase percentage of MVPA in PE	MVPA during PE; total daily PA	SOFIT SAPAC
Neumark-Sztainer (2003)	randomized controlled trial	n=201; 6 schools; mean age 15.4 (SD 1.1) years; females only	New Moves	•			•		•	16 w	positive changes in PA and dietary patterns	total daily PA; stage of PA behavioral change	self-report survey modified from Godin and Shepherd (1985)
Neumark-Sztainer (2010)	group randomized controlled trial	n=356; 12 schools; mean age 15.8 (SD 1.2) years; females only	New Moves	•			•		•	2 y	percentage body fat; BMI	daily PA; sedentary behaviors; stage of change PA behavior; PA goal setting; PA self-efficacy	3DPAR
Pate (2005)	group randomized controlled trial	n=2744; 24 schools; mean age 13.6 (SD 0.6) years ; females only	LEAP (Lifestyle Education for Activity Program)	•			•		•	1 y	percentage of girls reporting VPA and MVPA	total daily PA	3DPAR
Sallis (1997)	quasi-experimental	n=955; 7 schools; mean age range 9.49-9.62 years; both sexes	SPARK (Sports, Play, and Active Recreation for Kids)	•			•		•	2 y	PA levels in PE and out of school	total daily PA	accelerometers 1-day recall PA checklist SOFIT

				CSPAP components							Outcomes		
							SINV						
Study	Study design	Population	Intervention	QPE	PADSD	PABAS	SI	SW	FCE	Exp	Primary outcome	PA outcome	PA instrument
Sallis (2003)	randomized controlled trial	n=26,616 (approximately); 24 schools with mean enrollment 1109 (SD 356); 6th-8th grade; both sexes	M-SPAN (Middle-School PA and Nutrition)	•	•	•	•		•	2 y	PA levels at school	PA in PE; total PA at school; daily PA	SOFIT SOPLAY 7-day recall survey
Seo (2013)	pre-post no control intervention	n=1091; 8 schools; mean age 11.5 (SD 1.4) years; both sexes	HEROES (Healthy, Energetic, Ready, Outstanding, Enthusiastic, Schools) Initiative	•			•	•	•	18 mos	increase PA levels	total VPA and MPA	SHAQ
Springer (2012)	quasi-experimental (nonequivalent control)	n=511; 8 schools; mean age range 9.9 (SD 0.85)-10.0 (SD 0.80) years; both sexes	Marathon Kids		•		•		•	6 mos	PA engagement	time engaged in walking/running	AKP questionnaire PAC-Q
Webber (2008)	group randomized controlled trial	n=1721 (6th grade in 2003); mean age range 11.9-12.0 years; females only n=3504 (8th grade in 2005); mean age 14.0 years; females only n=3502 (8th grade in 2006); mean age 14.0 years; females only	TAAG (Trial of Activity for Adolescent Girls)	•			•		•	3 y	daily MET-weighted MVPA	total daily MVPA	accelerometers SOFIT

				CSPAP components							Outcomes		
							SINV						
Study	Study design	Population	Intervention	QPE	PADSD	PABAS	SI	SW	FCE	Exp	Primary outcome	PA outcome	PA instrument
Williamson (2007)	randomized field trial with two treatment arms	n=661; 4 schools; mean age 9.2 (SD 4.09) years; both sexes	Wise Mind Project		•		•		•	2 y	weight gain prevention	total daily PA	SAPAC
Young (2006)	randomized controlled trial	n=221; 1 school; 9th grade; females only	alternative PE class focused on life skills with a goal of increasing PA in PE and having family support	•			•		•	8 mos	daily energy expenditure; sedentary activities, cardiorespiratory fitness; cardiovascular disease risk factors	total daily PA	7-day PA Recall and self-reported sedentary activities questionnaire

Note: AKP, Active Kids Project questionnaire; BAS, before and after school; CATCH, Child and Adolescent Trial for Cardiovasacular Health; CATCH BP, Coordinated Approach To Child Health BasicPlus; CATCH BPC, CATCH BP and Community; CSH, Coordinated School Health approach; CSHP, Coordinated School Health Program; FCE, family and community engagement; MET-weighted MVPA, daily MET-weighted minutes of moderate-to-vigorous PA; mos, months; MPA, moderate PA; MVPA, moderate-to-vigorous PA; PA, PA; PADSD, PA during the school day; PAQ, PA Questionnaire; PDPAR, Previous Day PA Recall; QPE, quality physical education; SAPAC, Self-administered PA Checklist; SHAQ, Student Health Assessment Questionnaire; SD, standard deviation; SI, staff implementation; SINV, staff involvement; SOFIT, System for Observing Fitness Instruction Time; SOPLAY, System for Observing Play and Leisure Activity of Youth; SPAN, Student PA and Nutrition questionnaire; VPA, vigorous PA; w, weeks; y, years; YAQ, Youth Activity Questionnaire; year, academic year; 3DPAR, 3-day PA Recall; more detailed version of Table 3 available as supplemental document

• component present in intervention

[a] sub-sample for outcomes shown in this paper; [b] interim results from spring 2007-spring 2008

Table 3.4 Standardized mean difference random effects (Hedges's g) of Comprehensive School Physical Activity Promotion interventions on changes in youth total daily physical activity

	All Studies				Reporting of Physical Activity Outcome: Boys Only				Girls Only				Boys and Girls			
	n	*g*	(95CI)	I^2	*n*	*g*	(95CI)	I^2	*n*	*g*	(95CI)	I^2	*n*	*g*	(95CI)	I^2
Overall effect	15[a]	0.1	(0.03, 0.19)	90	5	0.1	(-0.10, 0.28)	92	10	0.1	(-0.02, 0.23)	90	5	0.1	(0.05, 0.19)	64
Sensitivity Analysis	13[b]	0.1	(0.02, 0.11)	65	4[c]	0	(-0.08, 0.11)		8[b]	0	(-0.06, 0.10)	66	4[d]	0.2	(0.10, 0.21)	
Number of CSPAP Components																
2	10	0.1	(-0.01, 0.14)	70	2	-0	(-0.34, 0.27)	85	7	0	(-0.09, 0.13)	71	3	0.1	(0.02, 0.20)	77
3	4	0.2	(0.07, 0.31)	93	2	0.2	(-0.12, 0.48)	91	2	0.2	(0.04, 0.43)	94	2	0.2	(0.01, 0.29)	0
4	1	0.3	(0.06, 0.53)		1	0.1	(-0.01, 0.28)		1	0.5	(0.17, 0.74)					
Specific CSPAP Component																
Quality Physical Education																
Y	12	0.1	(0.02, 0.19)	91	5	0.1	(-0.10, 0.28)	92	10	0.1	(-0.02, 0.23)	90	3	0.1	(0.01, 0.15)	57
Physical Activity During the School Day																
Y	5	0.2	(0.05, 0.32)	31	1	0.1	(-0.36, 0.63)		1	0.5	(0.10, 0.82)		4	0.2	(0.10, 0.21)	0
N	10	0.1	(-0.02, 0.17)	93	4	0.1	(-0.17, 0.32)	93	9	0.1	(-0.05, 0.19)	89	1	0	(-0.01, 0.08)	
Physical Activity Before/After School																
Y	1	0.3	(0.19, 0.40)		1	0.1	(-0.36, 0.63)		1	0.5	(0.10, 0.82)					
N	14	0.1	(0.02, 0.18)	90	4	0.1	(-0.17, 0.32)	93	9	0.1	(-0.05, 0.19)	89	5	0.1	(0.05, 0.19)	64
Staff Wellness																
Y	2	0.2	(0.04, 0.38)	97	2	0.2	(-0.08, 0.45)	56	2	0.2	(-0.02, 0.47)	94				
N	13	0.1	(0.02, 0.16)	77	3	0	(-0.20, 0.25)	95	8	0.1	(-0.05, 0.20)	86	5	0.1	(0.05, 0.19)	64
Family/Community Engagement																
Y	15	0.1	(0.03, 0.19)	90	5	0.1	(-0.10, 0.28)	92	10	0.1	(-0.02, 0.23)	90	5	0.1	(0.05, 0.19)	64
N																

Measure of Physical Activity																				
Objective	3	0	(-0.16,	0.20)	69	1	0	(-0.44,	0.44)		2	-0	(-0.33,	0.22)	89	1	0.1	(- 0.14,	0.39)	
Self-Report	14	0.1	(0.03,	0.20)	89	5	0.1	(-0.12,	0.28)	89	9	0.1	(-0.02,	0.26)	89	5	0.1	(0.05,	0.19)	63

[a] Total of 14 unique studies reporting 15 interventions. A single study, Sallis et al., 1997, reported outcomes for two interventions and are treated separately for the analyses; [b] Removal of Seo et al., 2013 and Sallis et al., 2003; [c] Removal of Seo et al., 2013; [d] Removal of Leupker et al., 1996

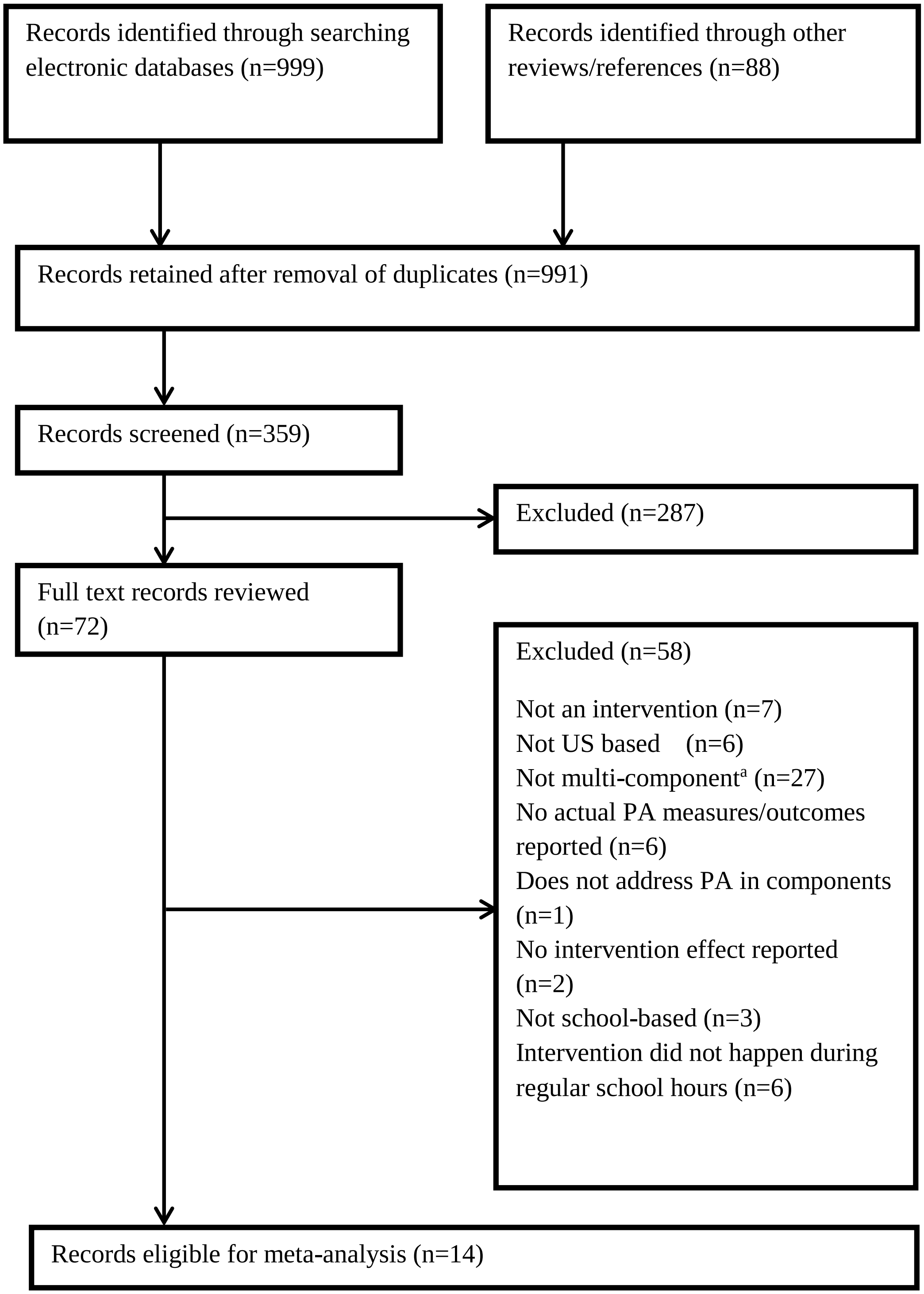

Figure 3.1. Flow chart of selection process resulting in inclusion of 14 unique records
[a] refer to Introduction for relevant components

Study name	Hedges's g	(95CI)
Caballero 2003	0.136	(0.041, 0.232)
Gortmaker 1999	0.004	(-0.107, 0. 116)
Hoelscher 2010	0.196	(-0.050, 0.443)
Leupker 1996	0.037	(-0.008, 0.081)
Neumark-Sztainer 2003	0.192	(-0.010, 0.394)
Neumark-Sztainer 2010	0.105	(0.001, 0.208)
Pate 2005	0.055	(-0.032, 0.142)
Sallis 1997[a]	-0.139	(-0.249, -0.029)
Sallis 1997[b]	-0.075	(-0.185, 0.035)
Sallis 2003	0.294	(0.190, 0.398)
Seo 2013	0.379	(0.319, 0.438)
Springer 2012	0.164	(0.064, 0.264)
Webber 2008	0.065	(-0.002, 0.133)
Williamson 2007	0.154	(0.061, 0.248)
Young 2006	0.117	(-0.153, 0.387)
Pooled Effect	0.110	(0.033, 0.187)

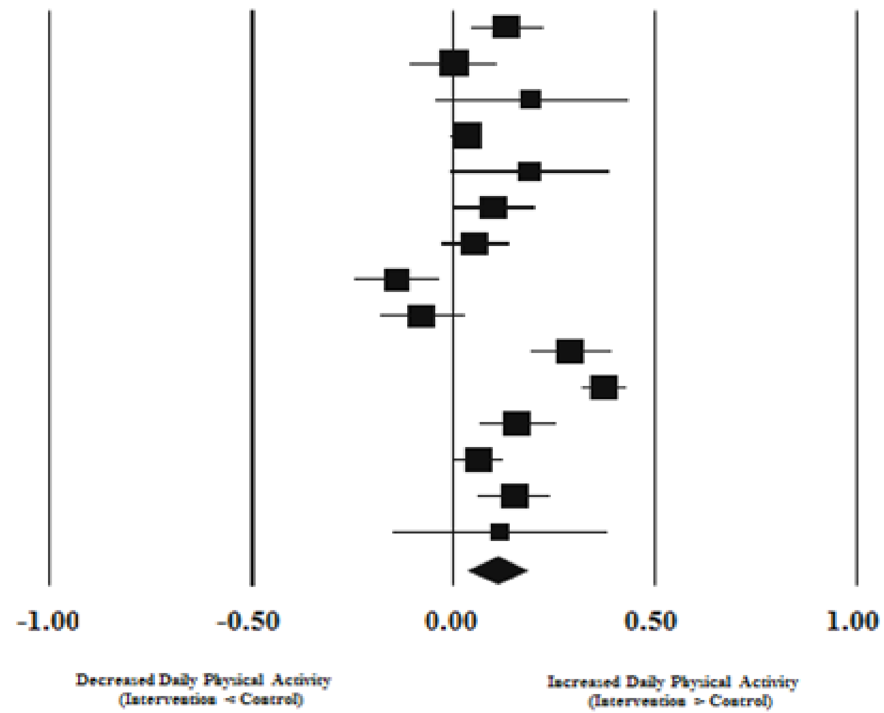

Figure 3.2. Forest plot of overall study standardized mean differences (Hedges's g) of Comprehensive School Physical Activity Promotion interventions on changes in youth total daily physical activity

CHAPTER 4: STUDY 2

DEVELOPMENT OF THE SYSTEM FOR OBSERVING STUDENT MOVEMENT DURING ACADEMIC ROUTINES AND TRANSITIONS (SOSMART)[1]

[1]Russ, L., Webster, C.A., Beets, M.W., Weaver, G., Egan, C.A., Harvey, R., & Phillips, D.S. To be submitted to *American Journal of Public Health.*

Physical activity (PA) is well documented as important and beneficial for children in many ways (CDC, 2013; CDC, 2010; IOM, 2013). Increasing PA is associated with improved health through reducing risk factors for diseases like obesity, Type 2 diabetes, and cardiovascular disease (CDC, 2013; McKenzie & Kahan, 2008). Being active is also associated with improvements to muscular strength, bone strength, self-esteem, and lower levels of anxiety and/or depression (CDC, 2013), thereby demonstrating the importance of PA to the mental and physical health of children (IOM, 2013). Further, increased amounts of PA during school have been associated with improved academic performance of children (CDC, 2010).

The U.S. Department of Health and Human Services (2008) recommends America's youth (6+ years old) engage in 60 minutes or more of moderate- to vigorous-intensity PA every day. Not only are children not meeting this recommendation (CDC, 2013; United States Department of Health and Human Services [USDHHS], 2008), but there is growing concern that children are increasingly sedentary, especially in their classrooms where they spend up to 9 hours each school day (CDC, 2013). While increasing opportunities for PA is important, reducing sedentary time may be equally important (IOM, 2013). Sedentary behaviors are associated with unfavorable health outcomes (Matthews et al., 2008) and may negatively affect children's health despite their engagement in PA (Biddle, Gorley, & Stensel, 2004; Dietz, 2001; Salmon, 2010).

Schools have been identified as a key setting to intervene (CDC, 2013; IOM, 2013; National Physical Activity Plan, 2012; Pate et al., 2006; USDHHS, 2008). Recommendations for increasing PA and reducing sedentary time include utilizing a multi-component approach through schools, including movement integration (MI) in the academic classroom (IOM, 2013; CDC, 2013). In elementary schools, the academic classroom is where generalist classroom teachers (CT) instruct students in academic subjects (e.g., math, language arts), and where students spend the majority of the school day. Integrating movement into the classroom setting has empirical support for making contributions to student PA (Bartholomew & Jowers, 2011; Beighle, Erwin, Beets, Morgan, & Le Masurier, 2010; Erwin, Beighle, Morgan, & Noland, 2011; Holt, Bartee, & Heelan, 2012; Mahar et al., 2006). Moreover, MI offers other benefits like decreasing sedentary time (Gortmaker et. al, 1999; Robinson, 1999; Salmon et al., 2005; Salmon, 2010), improving on-task behavior (Grieco, Jowers, & Bartholomew, 2009; Howie, 2013; Mahar et al., 2006; Mahar, 2011), increasing positive affect (Howie, Newman-Norlund, & Pate, 2014), and enhancing cognitive function (Donnelly & Lambourne, 2011; Elmakis, 2010; Howie et al., 2014).

Despite these benefits, little is known about the extent or nature of MI in schools (Webster, Russ, Vazou, Goh, & Erwin, 2015). Research on MI in non-intervention settings is scarce and has relied solely on teacher self-reports (Webster et al., 2013a; Elmakis, 2010; AAHPERD, 2011; Cothran, Kulinna, & Garn, 2010; Evenson, Ballard, Lee, & Ammerman, 2009; Holt et al., 2013). In the context of PA interventions through schools, the extent to which CTs are implementing MI as designed is also limited to self-reports (Bartholomew & Jowers, 2011; Cradock et al., 2014; Gibson et al., 2008; Howie et al, 2014; Kohl, Moore, Sutton, Kibbe, & Schneider, 2001; Kibbe et al., 2011; Skrade, 2013; Stewart, Dennison, Kohl, & Doyle, 2004; Williamson et al., 2007; Woods, 2011). One exception is the Move-To-Improve (MTI) classroom-based PE program (Dunn, Venturanza, Walsh, & Nonas, 2010). The primary objective of the MTI program was to help students meet the New York state requirement for PE minutes. Full-day classroom observations were conducted, which focused on MI strategies that were key areas of focus within the MTI intervention (i.e., frequency and duration of physical activities, teacher participation and/or encouragement, and academic content incorporated).

The extent and nature of MI across diverse classroom settings have not been objectively quantified through systematic observation. Systematic observation is a proven method of capturing contextual and behavioral variables that are useful in operationally defining, advancing, and evaluating best practices in teaching (Flanders, 1970; Flanders, 1976; van der Mars, 1989) and physical activity promotion in a number of settings. Examples include the System for Observing Fitness Instruction Time (SOFIT) in physical education (McKenzie, Sallis, & Nader, 1992), the System for Observing Play and Leisure Activity in Youth (SOPLAY) in school settings (McKenzie, Marshall, Sallis, & Conway, 2000), the System for Observing Play and Recreation in Communities (SOPARC) in community parks (McKenzie, Cohen, Sehgal, Williamson, & Golinelli, 2006), and the System for Observing for Staff Promotion of Activity and Nutrition in afterschool programs and summer day camps (Weaver, Beets, Webster, & Huberty, 2014).

The purpose of systematic observation is to provide a permanent record of events or activities that occurred to be analyzed at a future time and is typically used in research and supervision (van der Mars, 1989). An underlying assumption is that focusing only on events or behaviors that can be directly observed is believed to generate a more accurate account than self-reports. Major advantages of systematic observation for assessing PA include flexibility, low levels of inference, the ability to capture information about the physical and social environments at the same time, minimal interference with participants, and results that are easily quantifiable and often summarized in a way that is easy for policy makers, administrators, and practitioners, to understand (i.e. frequency, duration, percentage of total time) (McKenzie & van der Mars, 2015).

An instrument designed to systematically observe classroom-based strategies for increasing PA and reducing sedentary time can be used to measure implementation fidelity of MI interventions and provide empirical evidence of what transpires in the academic classroom context. This information is currently absent from the research literature on multicomponent efforts to increase youth PA through schools, which have been minimally effective (Russ, Webster, Beets, & Phillips, 2015). Providing such information would extend the descriptive knowledge base that informs policy decisions and program evaluation in the context of school wide efforts to promote PA. In addition, there is a burgeoning field of implementation science that acknowledges the need for examining the implementation and uptake of interventions. Evidence of increased interest in implementation science can be seen in the launching of the *Implementation Science* journal (Eccles & Mittman, 2006) and the NIH Dissemination and Implementation conference (Proctor et al., 2009), the appointing of special funds by the NIH reserved for grants explicitly studying dissemination and implementation, and emerging research examining the gap between research findings and practice (Damschroder, Aron, Keith, Kirsh, Alexander, & Lowery, 2009; Proctor et al., 2009). A systematic observation instrument designed to capture MI can be also be used for educational purposes. Such an instrument can yield information needed to enhance future recommendations for pre-service teacher education and in-service teacher training by translating findings into practical strategies for teachers to integrate movement in settings similar to their own.

Conceptual Framework

MI is defined as opportunities that allow for reduced sedentariness and/or increased PA among children during normal classroom time (Webster et al., 2015). MI encompasses the promotion of PA at any intensity (light, moderate, or vigorous; IOM, 2013). Current recommendations for MI focus on two major strategies: (a) incorporating PA breaks between academic lessons, and (b) infusing PA into academic lessons (Webster et al., 2015). PA breaks between lessons, also called exercise breaks (Elmakis, 2010) or PA breaks (CDC, 2013; IOM, 2013) are usually 10-15 minute sessions led by the CT, intended to require little planning or equipment (e.g. stretches, jogs around the classroom, jumping with an invisible rope, series of Yoga poses; Elmakis, 2010; CDC, 2013; IOM, 2013; Katz et al., 2010; Orlowski, Lorson, Lyon, & Minoughan, 2013). Other examples of PA breaks include Energizers (Mahar et al. 2006), chair aerobics (Ahamed et al., 2007), activity break cards (Erwin et al., 2011), and active transitions (Elliot, Erwin, Hall, & Heidorn, 2013; Orlowski & Hart, 2010).

Integrating PA into academic content can involve using an existing integrated PA curriculum (e.g. Move For Thought; Skrade & Vazou, 2013; SPARKabc's; www.sparkpe.org/abc/sparkabc/; Take 10!; Stewart et al., 2004), or combining existing lessons with an existing PA program, or modifying lessons to include an existing program (Bartholomew & Jowers, 2011; Donnelly et al., 2009; Grieco et al., 2009;). Also referred to as "content-rich" activities, these are lessons where PA is intentionally connected to a student learning objective (Erwin, Beighle, Carson, & Castelli, 2013; Castelli & Ward, 2012).

Purpose of the Study

The recent growth of the field of implementation science demonstrates the desire of researchers to examine the gap between findings and implementation. Measuring implementation fidelity may help explain the limited effectiveness of multi-component school-based PA interventions (Russ et al., 2015). Currently, however, objective measures for classroom-based strategies to increase PA and reduce sedentary time are limited. Given the advantages of systematic observation as an objective method for both research and practice related to PA promotion (McKenzie & van der Mars, 2015), the purpose of this study was to describe the development, reliability, and validity of a systematic observation instrument designed to measure MI. The instrument – named the System for Observing Student Movement during Academic Routines and Transitions (SOSMART) – will be useful in future research to determine the extent of MI, specifically to describe fidelity of MI intervention implementation, identify possible limitations in its use, and develop optimal strategies for increasing its effectiveness and sustainability as a key component of school-based PA promotion.

Methods

Participant Selection

Participants for this study included CTs (*N*=20, mean age=34.9 years, sd=10.4) and their students in existing, intact classes in grades 1-5 at four elementary schools in the Columbia, South Carolina area. The schools were selected based on their existing collaborative relationships with the research institution. The schools are situated in two different school districts (two schools from each district). The two schools in the first district served a combined total of approximately 964 students in grades K-5 with 58.6% of the students eligible for free and reduced lunch (South Carolina State Department of Education, 2013). The two schools in the second district served a combined total of approximately 376 students across grades K-3. Eligibility for free and reduced lunch data was not available for these schools at the time of the study.

Approval to conduct the study was obtained from the university IRB and from each school district. Informed consent was obtained from the teachers during an orientation meeting prior to sample selection. Purposeful sampling was used to ensure access to CTs demonstrating MI in and across diverse contexts (i.e. grade level and class size). This was achieved through administering a survey to all CTs, at all four schools, who provided consent to participate. The purpose of the survey was to identify classrooms that would be most useful in developing an instrument that would capture a variety of MI strategies and the frequency with which MI strategies are utilized. CTs responded to a self-report measure of PA promotion in the academic classroom (adapted from Webster, et al., 2013a) and demographic questions including teacher background variables (e.g. age, years of teaching experience, highest level of education) and classroom context variables (e.g., teacher-student ratio, socio-economic status of the students, grade level). The survey was developed and adapted with insight from previous research (AAHPERD, 2011; Elmakis, 2010; Webster, et al., 2013a), two MI scholars, and three CTs to ensure content validity. The survey data were used to identify the classrooms at each school with the highest prevalence and variety of MI strategies. The first step was to remove Pre-K, Kindergarten, Special Education, and specialized instructors (i.e., reading interventionists) from the sample/responses because we felt those contexts were more specialized situations and less representative of a general teacher's classroom. Responses from the remaining CTs were coded, categorized, and then sorted (within each subcategory) by grade level, number of students, number of assistants, content areas used for MI, frequency of MI, variance of MI, and the highest combined score for frequency and variety of PA promotion. Out of 80 survey respondents, 17 CTs were purposefully selected for the sample that provided representativeness across a variety of contextual variables (i.e. grade level, number of students) and provided the greatest likelihood of capturing a variety of MI strategies.

Scheduling conflicts and teacher dropout resulted in the need to identify seven additional participants. Therefore, two additional sampling strategies were employed. First, any CTs that were not previously selected for the original sample were contacted for inclusion in this study. Second, graduate students and researchers not involved with this study were asked for recommendations about CTs seen using MI at these schools. Teachers identified from this step were contacted for inclusion in this study.

Procedure for Instrument Development

Four phases were utilized to develop SOSMART and examine its reliability and validity: Phase I: Establishing an *A Priori* Framework; Phase II: Expanding and Refining *A Priori* Framework; Phase III: Devising a System for Coding and Interpretation, and Phase IV: Reliability and Validity Testing.

Phase I: Establishing an *A Priori* Framework. The purpose of Phase I was to develop a framework to guide initial observations and develop content validity. An extensive review of the literature concerning MI, including research and recommendations, was used to establish an *a priori* conceptual framework. The initial framework conceptualized MI as containing three categories of deliberate movement: morning movements, PA infused into academic lessons, and PA breaks between lessons. These deliberate opportunities indicated a PA opportunity directed by the teacher. This bears some similarity to the teacher behavior categories (e.g. Gives Information, Gives Directions) and student response category (e.g. Student Predictable Response) of the Cheffer's Adaptation of the Flanders Interaction Analysis System (CAFIAS) (Cheffers & Mancini, 1989). In this sense, the category Student Predictable Response captures when students participate in teacher-directed activities or obey teacher instructions.

However, there are also subtle ways CTs can integrate PA opportunities in the classroom. These opportunities may be considered incidental because the activity was *not* directed by the teacher at the moment it happens. This activity could be the result of some routine or procedure put in place earlier in the year. Again, there is some similarity between incidental MI and a student response category from the CAFIAS systematic observation tool. For example, the Student Initiative Behavior category captures behavior that is not teacher directed (Cheffers & Mancini, 1989). Examples of incidental opportunities may include a procedure requiring students to walk around the perimeter of the classroom each time they need to sharpen a pencil. Another strategy, informed by recommendations in the literature, that may facilitate incidental opportunities for movement is to arrange the classroom in a particular way (i.e. placement of desks) or converting normally fixed structures, like desks, to moveable structures, so objects can be rearranged quickly to facilitate movement (Erwin, 2009; IOM, 2013). Whether deliberate or incidental, MI results in an opportunity for students to not be sedentary, regardless of the intensity level of the movement. This means students can be engaged in light-, moderate-, or vigorous-intensity PA (IOM, 2013).

Phase II: Expanding and Refining *A Priori* Framework. The purpose of Phase II was to observe real-world examples of MI and determine if the *a priori* framework needed to be expanded and/or refined, and to further develop content validity through a Delphi survey. Trained researchers collected observational data by using one digital video camera to capture the classroom teacher and all students, when possible (with teacher and parent consent). The camera was operated using a tripod and set up unobtrusively in a corner of the classroom. Classroom observations occurred on regularly scheduled school days during normal classroom time with existing, intact classes. Across all classrooms, 32.4 total hours of videotaped observations were collected with an average observation time of 1.6 hours. Observations were conducted at times that did not overlap with state mandated testing times or occur during the first or last month of the school year. On each classroom visit, academic lessons and any transitions were recorded.

As data were collected, the lead researcher began viewing the videos to catalogue examples of MI. The *a priori* conceptual framework guided initial observations, although the researcher also remained sensitive to unanticipated MI behaviors or opportunities. Video examples and initial categories of MI were discussed with a second researcher whenever questionable behaviors or opportunities emerged. In such cases, if the identified behavior/opportunity was not readily catalogued using the *a priori* conceptual framework, the framework was revised (Webster et al., 2013b). Consistent with previous instrument development procedures, video viewings and discussions continued throughout data collection and afterward to confirm and expand MI concepts until the observations yielded no further insight (Thomas, Nelson, & Silverman, 2011; Weaver et al., 2014).

Following the development of initial MI concepts from the video data, a Delphi survey was utilized to confirm and/or expand these concepts and further develop content validity. Participants were provided with the definition of MI (Webster et al., 2015) and then asked to respond to an open-ended prompt (i.e. Classroom movement integration (MI) involves reducing your students' sedentary time (e.g., sitting) and/or increasing their physical activity during normal classroom time (i.e., in elementary general education classrooms). Please list all examples and/or strategies you can think of that represent MI.) The survey was sent electronically to individuals identified as experts in the field. Experts were classified as (a) scholars in higher education with experience teaching and/or researching MI, or (b) practicing classroom teachers in the elementary school setting. Eighty-five experts (46 scholars in higher education/research and 39 practicing classroom teachers) were contacted via e-mail with a request for participation. The first round was exploratory in nature (Thomas et al., 2011; Weaver et al., 2014). Thirty-two responses (12 scholars and 20 teachers) were received, providing a 38% response rate. Delphi responses were used to confirm and expand the categories. Then, a second round was sent out to all respondents for additional feedback. The second round yielded no further insights; therefore, no further rounds were pursued.

The final MI concepts and their operational definitions are presented in Table 4.1. The instrument uses a two-stage decision-making process focused first on teacher involvement and then on student responses. Teacher involvement is described by three categories: the person giving the directive to be active (i.e. classroom teacher or other), instructional variables (i.e. the teacher led the activity or technology was used to lead the activity), and movement type variables (i.e. deliberate MI as a reward/incentive, opening activity, transition, and/or other movement that was academic or non-academic in nature). Student involvement is described by two categories: the part of the class that was active (i.e. whole class, part class, or small group) and the reason for it (i.e. in response to the deliberate teacher directive, or incidentally as a result of the physical environment or a non-teacher directed transition).

Phase III: Devising a System for Coding and Interpretation. The purpose of Phase III was to create a coding scheme and strategy for summarizing and/or interpreting the instrument results. SOSMART was designed to be an interval recording system to capture the variety and frequency of MI opportunities, which are theorized to lead to physically active student responses. Inactive vs. active are operationally defined as follows:

- Inactive- student(s) engaged in sedentary or low-active behaviors (i.e. lying down, sitting, standing quietly (Marshall & Merchant, 2013; McKenzie et al., 2002; Weaver et al., 2014; Welk, 2002).
    - *Note:* This excludes standing and stretching (i.e. performing non-locomotor movements while sitting and/or standing. These behaviors are included in “active” (see below).

- Active- student(s) engaged in locomotor movement (ranging from walking to running) and/or isolated upper body and/or lower body movements (non-locomotor) whether sitting or standing.
    - *Note:* Using these definitions, sitting on an exercise ball is not sitting at rest. Therefore, it is active.

Coding Procedure. For each interval, decisions must be made about teacher involvement and student response. The first stage requires a decision to be made about the involvement of the classroom teacher by answering the following question: Did the classroom teacher give a direction to be active? If the answer is Yes, the observer moves on to code teacher involvement behaviors (teacher directive variables, instruction variables, and movement variables), then proceeds to Stage 2 (student response variables). If the answer is No, the observer moves on directly to code Stage 2 (student response variables).

The second stage requires a decision to be made about the response of the class by answering the following question: How did students respond? If the answer to the previous stage was Yes, the observer records what part of the class is active (whole class, part class, or small group). Context variables identify how much of their body is active (upper body only, lower body only, or full body) and off-task behavior. If the answer to the previous stage was No, the observer records what part, if any, of the class is active and the observable reason for that movement (as a result of something in the physical environment or as a result of a non-teacher directed transition, like getting supplies or using the bathroom). Within these categories, context variables identify the presence of added activity and/or off-task behavior. A flow chart illustrating the two stage decision making process is presented in Figure 1.

On prepared coding forms (Figure 2), trained observers list all relevant codes during continuous observation for 20-second intervals. When coding, the observer should list the appropriate code(s) in the appropriate 20-second cell as soon as evidence is observed. The observer should only list the code once in a given 20-second cell on the coding form, even if it is observed more than once during that interval. Context codes should be written as a sub-script to the major variable code. Coding a (-) is acceptable for consecutive cells when the movement continues across multiple consecutive intervals.

Interpretation Procedure. SOSMART is designed to capture observable MI variables and translate findings into an easily quantifiable format. The summary sheet (Figure 3) provides space to calculate the total number of intervals for each category. Total percentage of occurrence can be calculated as: $\text{Percentage occurrence} = \frac{\text{total number category intervals}}{\text{total number intervals in observation}} \text{x}100.$

A percentage of occurrences can be calculated for each code, as well as a tally mark for each unique instance of the code. There is no benchmark for high MI versus low MI frequencies or percentages of total time. Instead, SOSMART should be used to document the frequency and variety of MI strategies employed by teachers in the classroom. Continued research with this instrument may provide a better picture of what an appropriate benchmark might be for MI in the classroom setting (Webster et al., 2013b).

Phase IV: Reliability and Validity testing. The purpose of Phase IV was a) to test inter and intrarater reliability of the instrument, b) to further examine content validity, and c) to test construct validity of the instrument.

Observer Training and SOSMART Reliability. Consistent with previous research (Pope, Coleman, Gonzalez, Barron, and Heath, 2002) and recommendations (McKenzie and van der Mars, 2015), reliability training and testing followed a specific sequence of steps (i.e. orientation to systematic observation and the SOSMART instrument, committing behavior categories/codes to memory, video practice, live practice, and formal reliability) and consisted of three sessions. The first session was video practice, including booster training sessions, the second session was live practice, and the third session was used for reliability. Reliability was established through interobserver reliability and intraobserver reliability. Five observers not directly involved in instrument development (Phase II) were trained to use the instrument using video samples over a week long time period that included formal training by the primary author followed by a mid-week booster training. Training and observations occurred until 80% interobserver agreement was reached (Weaver et al., 2014). Two observers conducted field reliability live and two different observers conducted reliability from the same observation viewed on video.

SOSMART Validity. Two validity procedures were used in this phase. A Delphi survey was used to further examine content validity by identifying initial MI categories from the literature and recommendations, then considering those categories in light of direct observation of classroom teachers, and finally through reaching consensus of MI categories with experts. Statistical analysis was used to test the hypothesis that the presence of MI variables (teacher directives, instructional, and movement types) would contribute to student activity and/or decrease student inactivity. Construct validity of the instrument was evaluated by examining the presence/absence of teacher MI compared with students' activity and/or sedentary behaviors as measured with accelerometers from a sub-sample of 12 observations. The majority of these observations (n=10) were randomly selected within and across each grade level at each school to provide a representative picture across all four schools. In addition to random selection, additional observation (n=2) were purposefully selected for testing construct validity because they provided the greatest likelihood of seeing a variety of MI concepts.

Data Analysis. Statistical analyses were completed using STATA (v. 13.0, College Station, TX). Reliability for SOSMART was calculated using interobserver reliability and intraobserver reliability. Interobserver reliability (IOR) was measured by calculating interval-by-interval percent agreement as $\mathrm{IOR} = \frac{\text{agreements}}{\text{agreements+disagreements}}\mathrm{x}100$ (Mahar, 2011; Weaver et al., 2014; Webster et al., 2013b). Intraobserver reliability was measured using the test-retest (different day) method across a two-week span to examine the consistency of SOSMART across different days (Thomas et al., 2011; Webster et al., 2013b). Interval-by-interval percent agreement was calculated the same way. Validity of SOSMART was conducted by examining the presence/absence of MI variables compared to the activity counts per minute from the accelerometers using unconditional multilevel random effects logistical regression (Guo & Zhao, 2000). The choice was made not to separate boys and girls in analyses. Based on recent research (Bailey et al., 2012) and results from the Delphi survey, there was no reason to believe there would be a difference in activity between genders in the classroom. Separate models were estimated for each of the eleven MI variables. A cut-point of 100 counts/min was used (Matthews et al., 2008), where greater than 100 counts/min was considered active (i.e. total activity, regardless of intensity) and 100 counts/min or less was considered inactive.

Results

Reliability

IOR agreement and total reliability exceeded 80% in live and video reliability testing (Table 4.1). Intraobserver agreement across two weeks resulted in 97.5% agreement. Three MI variables were not observed (i.e. reward, other movement (academic), physical environment); therefore, reliability was not calculated for these variables.

Validity

Logistical regression models of MI variables related to total activity (i.e. activity counts/min) are presented in Table 4.2. Results support the hypothesis that students were more likely to be active when MI variables were present with 8 out of 11 variables achieving statistical significance (see Table 4.2). The strongest predictor of student activity was the presence of “other movement, academically infused”, suggesting that students are more likely to be active when MI that included teaching or reviewing academic content is present (Figure 4.4). The purpose of Figures 4.4 - 4.6 is to visually represent a sample demonstrating construct validity. That is, when MI is coded, student activity is more likely to be present. This data was purposefully selected from a teacher demonstrating the greatest frequency of MI implementation and variety of MI strategies during observations in order to provide the greatest number of examples illustrating construct validity. The activity data is from a randomly selected student within the class. Figure 4.4 is a graphical representation illustrating construct validity for the two strongest MI variables (e.g. OM (a) and OM (na)). This student is more likely to be engaged in total activity when the variables “other movement, academic” and/or “other movement, non-academic” are present. This student was also more likely to be active when teacher-directed transitions were present (Figure 4.5), especially when those transitions were deliberately infused with PA (TT+). As expected, when a teacher directive to be active occurred, this student was more likely to be in activity; similarly, in the absence of a teacher directive, this student was not active (i.e. registered <100 counts/min on the accelerometer).

What is interesting about these illustrations are the different responses to different MI variables (i.e. the activity peak for OM(a) is higher than the peak for OM(na), Figure 4.4), and the presence of activity (i.e. peaks of activity counts) in the absence of any MI variables. A possible reason for seeing a greater peak in activity for OM(a) as compared to OM(na) may be that this particular student is more interested in, or more motivated by, activities where academic content is incorporated into the movement. Thus, it is possible that the difference in student response between these two MI variables depends on characteristics of the student.

In relation to teacher-directed transitions (Figure 4.5), there are moments when a teacher-directed transition is present; however, this particular student is minimally active. This may be an instance where the teacher is releasing students to or from a location by small groups (i.e. releasing one table or pod at a time to line up for lunch) and this student’s group was simply not called yet. Other instances where the student is active in the absence of any MI variables (Figure 4.5 and Figure 4.6), may be illustrations of incidental movements (i.e. going to the bathroom or getting supplies) that are not deliberately directed by the teacher. These moments (NT, Figure 4.6), are an indication that something else was facilitating activity. It may have been a non-teacher directed transition (i.e. getting a supply or housekeeping tasks like going to the bathroom); or, it may have been something in the environment that was facilitating activity. In Figure 4.6, something in the environment (i.e. a fit stool) was facilitating the movement during the non-teacher directed transition and may be considered an example of incidental MI.

Despite using the established literature and the Delphi survey to content validate all of the SOSMART variables, we were not able to demonstrate construct validity with statistical significance for three of the variables (reward, opening activity, physical environment).

Discussion and Conclusion

To our knowledge, SOSMART is one of the first systematic observation tools for measuring the frequency and variety of MI strategies utilized in the academic classroom. This instrument fills the need for objective measurements of MI in the academic classroom setting, which is included as a key context in coordinated and comprehensive approaches to PA promotion through schools (CDC, 2013; IOM, 2013). While SOSMART was found to be valid and reliable overall, three MI variables were not observed enough to establish construct validity. In terms of their validity, these variables were present less frequently than the other eight variables. It may be that these variables are referred to less frequently, if at all, in the literature and current MI recommendations.

More discussion of how to incorporate these MI strategies in practice may be needed, and continued use of SOSMART is needed to further validate these variables.

The figures (Figure 4.4-4.6) not only illustrate differences between activity peaks, but also peaks and valleys where we may or may not expect them. While the purpose of this study was not to understand these differences, future research should examine these differences and explore reasons underpinning the presence of them. It is possible that different MI variables, or combinations thereof, can have different activity outcomes. For example, different strategies may be more or less effective depending on any number of student variables (i.e. student interest, attitude, experience, or even the actual number of students in the class). Therefore, documenting these differences and exploring the underlying reasons for them has implications for practice. Specific MI strategies may or may not be recommended to preservice and/or inservice CTs depending on their school or classroom context. This instrument also provides MI terms that can be used as a common language in communicating about MI during preservice teacher training and inservice teacher development.

Even though the figures represent a high promoting teacher, and a randomly selected student, these illustrations may not represent all cases. Therefore, descriptive research is needed to provide a more comprehensive picture of how MI is being used in and across a variety of classrooms. The data obtained from SOSMART will also enable researchers to evaluate intervention implementation fidelity. Descriptive research and implementation science can contribute to component-specific national surveillance data needed to strengthen the effectiveness of CSPAP efforts. This will not only benefit evaluations of program effectiveness, but may also be used in policy and practice decisions. For example, MI research can fuel efforts to establish a benchmark policy, or national recommendation, for MI in the classroom setting. MI may also be given consideration by school administrators in the practice of annual evaluations of CTs.

It must also be acknowledged that the data generated from using this instrument provide descriptive, but not prescriptive, information (McKenzie & van der Mars, 2015). Researchers are cautioned to remember that systematic observation findings are always contextual and limited due to human error (van der Mars, 1989). Common sources of observer error include observer drift, reactivity, environmental factors, and bias or falsifying data (McKenzie & van der Mars, 2015). These should be addressed and carefully safeguarded against throughout training and data collection.

Future research directions should include using SOSMART to provide a descriptive knowledge base about the extent and nature of MI, examining which MI variables are more/less feasible in certain classroom contexts (e.g., with larger vs. smaller class sizes), and using SOSMART to evaluate implementation fidelity in classroom-based PA interventions. SOSMART can also be used in combination with other systematic observation measures (i.e. SOFIT in physical education) to improve surveillance research on CSPAP prevalence. To our knowledge, there currently is not an evidence-based benchmark for the amount of MI that should be implemented in the classroom context. Recommendations for increasing student activity and/or decreasing sedentarism in the classroom could be revised using a stronger empirical basis.

References

Ahamed, Y., Macdonald, H., Reed, K., Naylor, P. J., Liu-Ambrose, T., & Mckay, H. (2007). School-based physical activity does not compromise children's academic performance. *Medicine and Science in Sports and Exercise, 39*(2), 371.

American Alliance for Health, Physical Education, Recreation, and Dance [AAHPERD]. (2011). 2011 Comprehensive School Physical Activity Survey.

Bailey, D. P., Fairclough, S. J., Savory, L. A., Denton, S. J., Pang, D., Deane, C. S., & Kerr, C. J. (2012). Accelerometry-assessed sedentary behaviour and physical activity levels during the segmented school day in 10–14-year-old children: the HAPPY study. *European Journal of Pediatrics*, *171*(12), 1805-1813.

Bartholomew, J. B., & Jowers, E. M. (2011). Physically active academic lessons in elementary children. *Preventive Medicine, 52*, S51-S54.

Beighle, A., Erwin, H. E., Beets, M. W., Morgan, C. F., & Le Masurier, G. (2010). America on the move: School-based physical activity promotion. *International Journal of Physical Education, 47*(2), 2-16.

Biddle, S. J., Gorely, T., & Stensel, D. J. (2004). Health-enhancing physical activity and sedentary behaviour in children and adolescents. *Journal of Sports Sciences*, *22*(8), 679-701.

Castelli, D. M., & Ward, K. (2012). Physical activity during the school day. *Journal of Physical Education, Recreation & Dance*, *83*(6), 20-29.

Centers for Disease Control and Prevention [CDC]. (2013). Comprehensive School Physical Activity Programs: A guide for schools. Retrieved from http://www.cdc.gov/healthyyouth/physicalactivity/cspap.htm.

Cheffers, J. & Mancini, V. (1989). Cheffers' Adaptation of the Flanders' Interaction Analysis System (CAFIAS). In Darst, P., Zakrajsek, D., & Mancini, V. (eds.), *Analyzing Physical Education and Sport Instruction.* Champaign, IL: Human Kinetics.

Cothran, D. J., Kulinna, P. H., & Garn, A. C. (2010). Classroom teachers and physical activity integration. *Teaching and Teacher Education*, *26*(7), 1381-1388.

Cradock, A., Barrett, J., Carter, J., McHugh, A., Sproul, J., Russo, E., … & Gortmaker, S. (2014). Impact of the Boston Active School Day policy to promote physical activity among children. *American Journal of Health Promotion, 28*(3) S54-S64.

Damschroder, L. J., Aron, D. C., Keith, R. E., Kirsh, S. R., Alexander, J. A., & Lowery, J. C. (2009). Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. *Implementation Science, 4*(1), 50.

Dietz, W. H. (2001). The obesity epidemic in young children. *British Medical Journal, 322*(7282), 313-314.

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.

Donnelly, J. E., & Lambourne, K. (2011). Classroom-based physical activity, cognition, and academic achievement. *Preventive Medicine*, *52*, S36-S42.

Dunn, L. L., Venturanza, J. A., Walsh, R. J., & Nonas, C. A. (2012). An observational evaluation of Move-To-Improve, a classroom-based physical activity program, New York City Schools, 2010. *Preventing Chronic Disease, 9.*

Eccles, M. P., & Mittman, B. S. (2006). Welcome to implementation science. *Implementation Science, 1*(1), 1-3.

Elliot, E., Erwin, H., Hall, T., & Heidorn, B. (2013) Comprehensive school physical activity programs: Helping all students achieve 60 minutes of physical activity each day. *Journal of Physical Education, Recreation & Dance*, *84*(9), 9-15

Elmakis, G. S. (2010). *Survey of physical activity in elementary school classrooms in the state of Virginia.* (Doctoral dissertation). The College of William and Mary, Virginia.

Erwin, H. E. (2009). Implementing classroom-based physical activity. *Research Quarterly for Exercise and Sport,* 80(1), A10-A11.

Erwin, H. E., Beighle, A., Morgan, C. F., & Noland, M. (2011). Effect of a low-cost, teacher-directed classroom intervention on elementary students' physical activity. *Journal of School Health, 81*(8), 455-461.

Erwin, H., Beighle, A., Carson, R. L., & Castelli, D. M. (2013). Comprehensive school-based physical activity promotion: A review. *Quest*, *65*(4), 412-428.

Evenson, K. R., Ballard, K., Lee, G., & Ammerman, A. (2009). Implementation of a School-Based State Policy to Increase Physical Activity. *Journal of School Health*, *79*(5), 231-238.

Flanders, N. A. (1970). *Analyzing teacher behavior*. Addison-Wesley P. C.

Flanders, N. A. (1976). Interaction analysis and clinical supervision. *Journal of Research & Development in Education*.

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. *International Journal of Behavioral Nutrition and Physical Activity*, *5*(1), 36.

Gortmaker, S. L., Peterson, K., Wiecha, J., Sobol, A. M., Dixit, S., Fox, M. K., & Laird, N. (1999). Reducing obesity via a school-based interdisciplinary intervention among youth: Planet Health. *Archives of Pediatrics & Adolescent Medicine*, *153*(4), 409-418.

Grieco, L. A., Jowers, E., & Bartholomew, J. B. (2009). Physically active academic lessons and time on task: the moderating effect of body mass index. *Medicine and Science in Sports and Exercise, 41*(10), 1921-6.

Guo, G., & Zhao, H. (2000). Multilevel modeling for binary data. *Annual review of sociology*, 441-462.

Holt, E., Bartee, T., & Heelan, K. (2013). Evaluation of a policy to integrate physical activity into the school day. *Journal of Physical Activity & Health*, *10*(4).

Howie, E. K. (2013). *Classroom exercise breaks and educational outcomes in elementary school students.* (Doctoral dissertation). University of South Carolina, South Carolina.

Howie, E. K., Newman-Norlund, R. D., & Pate, R. R. (2014). Smiles count but minutes matter: responses to classroom exercise breaks. *American Journal of Health Behavior*, *38*(5), 681-689.

Institute of Medicine [IOM]. (2013). *Educating the student body: Taking physical activity and physical education to school*. Washington, DC: The National Academies Press.

Katz, D. L., Cushman, D., Reynolds, J., Njike, V., Treu, J. A., Katz, C., …, & Smith, E. (2010). Putting physical activity where it fits in the school day: Preliminary results of the ABC (Activity Bursts in the Classroom) for fitness program. *Preventing Chronic Disease, 7*(4).

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*, S43-S50.

Kohl, H. W., Moore, B. M., Sutton, A. W., Kibbe, D. L., & Schneider, D. C. (2001). A curriculum-integrated classroom physical activity promotion tool for elementary schools: Teacher evaluation of Take 10!™. *Medicine & Science in Sports & Exercise*, *33*(5), S179.

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.

Mahar, M. T. (2011). Impact of short bouts of physical activity on attention-to-task in elementary school children. *Preventive Medicine*, *52,* S60-S64.

Matthews, C. E., Chen, K. Y., Freedson, P. S., Buchowski, M. S., Beech, B. M., Pate, R. R., & Troiano, R. P. (2008). Amount of time spent in sedentary behaviors in the United States, 2003–2004. *American Journal of Epidemiology*, *167*(7), 875-881.

McKenzie, T. L., Cohen, D. A., Sehgal, A., Williamson, S., & Golinelli, D. (2006). System for Observing Play and Recreation in Communities (SOPARC): reliability and feasibility measures. *Journal of Physical Activity & Health*, *3*, S208.

McKenzie, T. L., Marshall, S. J., Sallis, J. F., & Conway, T. L. (2000). Leisure-time physical activity in school environments: an observational study using SOPLAY. *Preventive Medicine*, *30*(1), 70-77.

McKenzie, T. L., Sallis, J. F., & Nader, P. R. (1992). SOFIT: System for observing fitness instruction time. *Journal of Teaching in Physical Education*, *11*(2).

McKenzie, T. L. & van der Mars, H. (2015). Top 10 Research Questions Related to Assessing Physical Activity and Its Contexts Using Systematic Observation. *Research Quarterly for Exercise and Sport, 86,* 13-29.

National Association for Sport and Physical Education [NASPE] (2013). *Comprehensive school physical activity programs: Helping all students achieve 60 minutes of physical activity each day* [Position statement]. Reston, VA: Author. Retrieved from http://www.shapeamerica.org/advocacy/positionstatements/pa/loader.cfm?csModule=security/getfile&pageid=4726

National Physical Activity Plan Alliance. (2012). National Physical Activity Plan.

Orlowski, M. A., & Hart, A. (2010). Go! Including Movement during Routines and Transitions. *Young Children, 65*(5), 88-93.

Orlowski, M., Lorson, K., Lyon, A., & Minoughan, S. (2013). My classroom physical activity pyramid: A tool for integrating movement into the classroom. *Journal of Physical Education, Recreation, & Dance, 84*(9), 47-51.

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

Pope, R. P., Coleman, K. J., Gonzalez, E. C., Barron, F., & Heath, E. M. (2002). Validity of a revised system for observing fitness instruction time (SOFIT). *Pediatric Exercise Science*, *14*(2), 135-146.

Proctor, E. K., Landsverk, J., Aarons, G., Chambers, D., Glisson, C., & Mittman, B. (2009). Implementation research in mental health services: An emerging science with conceptual, methodological, and training challenges. *Administration and Policy in Mental Health and Mental Health Services Research*, *36*(1), 24-34.

Robinson, T. N. (1999). Reducing children's television viewing to prevent obesity: a randomized controlled trial. *Journal of the American Medical Association*, *282*(16), 1561-1567.

Russ, L., Webster, C.A., Beets, M.W., & Phillips, D.S. (in press). Systematic review and meta-analysis of multicomponent interventions through schools to increase physical activity. *Journal of Physical Activity and Health.*

Salmon, J. O., Ball, K., Crawford, D., Booth, M., Telford, A., Hume, C., ... & Worsley, A. (2005). Reducing sedentary behaviour and increasing physical activity among 10-year-old children: overview and process evaluation of the ‘Switch-Play’intervention. *Health Promotion International*, *20*(1), 7-17.

Salmon, J. (2010). Novel strategies to promote children's physical activities and reduce sedentary behavior. *Journal of Physical Activity & Health*, *7*(3), S299.

Skrade, M. (2013). *Integrated classroom physical activity: Examining perceived need satisfaction and academic performance in children*. (Doctoral dissertation). Iowa State University, Iowa.

Skrade, M. & Vazou, S. (2013). Integrated classroom physical activity: Examining perceived need satisfaction and academic performance in children. *Journal of Sport & Exercise Psychology,* 35, S8-S12.

South Carolina Department of Education (2013). *E-rate Free and Reduced Meal Eligibility Data.* Retrieved from https://ed.sc.gov/data/erate/index.cfm

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.

Thomas, J. R., Nelson, J. K., & Silverman, S. J. (2011). *Research methods in physical activity*. Champaign, IL: Human Kinetics.

United States Department of Health and Human Services [USDHHS]. (2008). *Physical Activity Guidelines for Americans.* Retrieved from http://www.health.gov/paguidelines

van der Mars, H. (1989). Systematic observation: An introduction. In Darst, P., Zakrajsek, D., & Mancini, V. (eds.), *Analyzing Physical Education and Sport Instruction.* Champaign, IL: Human Kinetics.

Weaver, R. G., Beets, M. W., Webster, C., & Huberty, J. L. (2014). System for observing staff promotion of activity and nutrition (SOSPAN). *Journal of Physical Activity & Health*, *11*(1).

Webster, C. A., Caputi, P., Perreault, M., Doan, R., Weaver, G., & Doutis, P. (2013a). Elementary classroom teachers' adoption of physical activity promotion in the context of a statewide policy: An innovation diffusion and socio-ecologic perspective. *Journal of Teaching in Physical Education*, *32*(4), 419-440.

Webster, C., Russ, L., Vazou, S., Goh, T., & Erwin, H. (2015). Integrating Movement in Academic Classrooms: Understanding, Applying, and Advancing the Knowledge Base. *Obesity Reviews*

Webster, C. A., Wellborn, B., Hunt, K., LaFleche, M., Cribbs, J., & Lineberger, B. (2013b). MPOWER: An observation system for assessing coach autonomy support in high school varsity boys' soccer practices. *International Journal of Sports Science and Coaching*, *8*(4), 741-754.

Williamson, D. A., Copeland, A. L., Anton, S. D., Champagne, C., Han, H., Lewis, L., ... & Ryan, D. (2007). Wise Mind Project: A school-based environmental approach for preventing weight gain in children. *Obesity*, *15*(4), 906-917.

Woods, C. D. (2011). *Evaluation of Instant Recess® exercise breaks as a means for implementing LAUSD physical activity policy in elementary schools.* (Doctoral dissertation). University of California, Los Angeles.

Table 4.1 Operational Definitions of the SOSMART Instrument and Interrater Percent Agreement

		Interrater Reliability	
Variable	Operational definition	Percent agreement live	Percent agreement video
Teacher Involvement (TI)			
Teacher Direct		93.72	99.26
Classroom teacher	Teacher gave an explicit direction for students to be active.	89.87	98.73
No	No teacher direction for students to be active occurred.	95.31	99.47
Instruction		89.87	93.67
Teacher-led	The teacher led the activity.	88.75	93.67
Technology-led	The teacher used technology (i.e. YouTube videos, electronic media like GoNoodle or JustDance) to lead the activity. The adult did NOT actually lead the activity.	-	100.00
Movement Type		88.61	95.34
Reward/Incentive	Movement provided by the teacher as an obvious (explicitly stated) reward for providing a correct response or behavior in class.	-	-
Opening activity	Movement directed by the teacher **within the first 10 minutes of the official start of the school day**, followed by a class response resulting in student activity. (This may include a school-wide morning exercise on the news show, etc…)	-	100.00
Teacher-directed transition	The teacher gave a direction for students to be active resulting in students moving from point A to point B (i.e. desks to carpet) or between finishing one task and getting ready for next task (i.e. putting away supplies and/or transitioning from one instructional content to another instructional	87.03	96.66

	content). This includes housekeeping tasks and procedures (picking up/putting away supplies (pencils/paper, tissues, snacks), using restroom) when the teacher has students walk from point A to point B.		
Other Movement	**Non-academic (na):** Movement directed by the teacher *within a lesson* or *between lessons*, followed by a class response resulting in student activity that **does NOT include academic content** (often called "brain breaks" or "exercise breaks").	92.00	92.31
	Academic-infused (a): Movement directed by the teacher *within a lesson* or *between lessons*, followed by a class response resulting in student activity **that DOES review/teach academic content.**	-	-
Student Response (SR)			
Students active	The amount of students in the class that are active, as defined herein, at first glance	91.00	88.14
Whole class	All students are active.	70.58	92.10
Part class	More than 50% but less than all students are active.	56.25	80.00
Small group	Less than 50% of students are active.	92.30	88.71
None	No students are active.	97.81	91.91
As a result of		84.21	80.88
Physical environment	Equipment used that is facilitative of movement, resulting in student activity, regardless of level of intensity.	-	-
Non-teacher directed transition	The teacher did not give a direction for student(s) to be active, but the student(s) still engaged in physical activity.	84.21	80.88

	This includes when students walk from point A to point B for tasks that are not directed by the teacher (i.e. getting supplies, going to the teacher’s desk, going to the trash can, etc…).		
Across variables for all intervals		**91.32**	**91.94**

Note. “-“ indicates the behavior was never observed therefore percent agreement was not calculated

Table 4.2 Construct Validity of the SOSMART Instrument

	Total Activity		
	OR	p-value	(95% CI)
Teacher			
Classroom teacher	**1.5**	<0.0	(1.4, 1.6)
Other	**2.0**	<0.0	(1.4, 2.8)
Instruction			
Teacher-led	**1.5**	<0.0	(1.4, 1.6)
Technology-led	**1.6**	0.02	(1.1, 2.4)
Movement type			
Reward	4.8	0.1	(0.6, 38.7)
Opening activity[a]	-	-	-
Teacher directed transition	**1.3**	<0.0	(1.2, 1.5)
Other movement (non-academic)	**1.9**	<0.0	(1.6, 2.3)
Other movement (academic)	**2.3**	<0.0	(1.5, 3.5)
Resulting from environment	1.0	0.93	(0.7, 1.5)
Non-teacher directed transition	**1.2**	<0.0	(1.1, 1.3)

[a] Too few observations to estimate
Note. Statistically significant relationships are bolded.
Abbreviations: OR, odds ratio

SOSMART Observational System

Coding Protocol Flow Chart

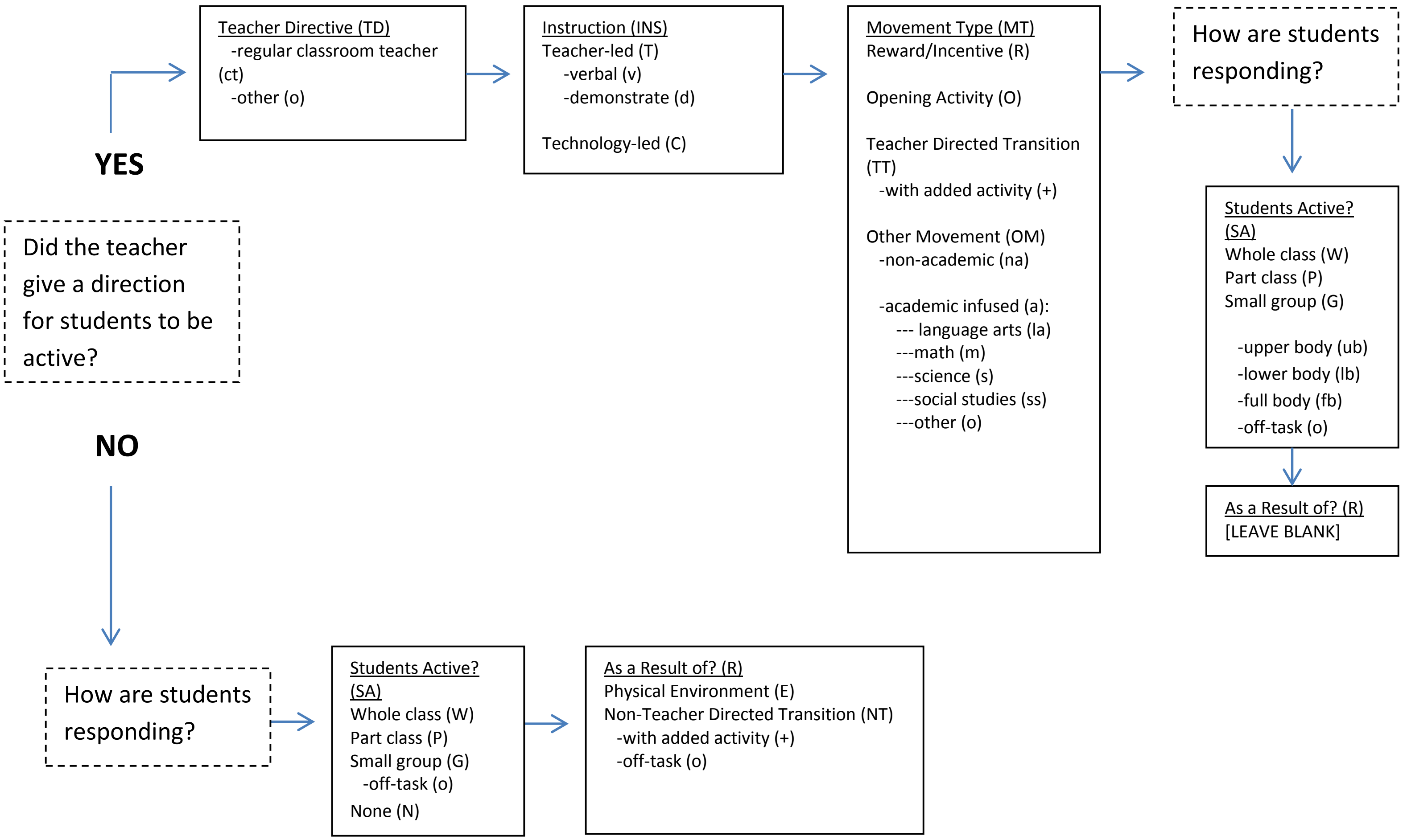

Figure 4.1 SOSMART decision flow chart

SOSMART Recording Sheet

School: ____________________________
Teacher Name:_____________________
Grade:_____________________________

Students: _______________________
Assistants:_______________________
Class time:
______AM/PM to _______AM/PM

Observer:________________________________
Observation Date:________________________
Coding start: ______AM/PM
Coding stop: ______AM/PM

		Intervals								
		1	2	3	4	5	6	7	8	9
TI	TD									
	INS									
	MT									
SR	SA									
	R									

		Intervals								
		1	2	3	4	5	6	7	8	9
TI	TD									
	INS									
	MT									
SR	SA									
	R									

		Intervals								
		1	2	3	4	5	6	7	8	9
TI	TD									
	INS									
	MT									
SR	SA									
	R									

		Intervals								
		1	2	3	4	5	6	7	8	9
TI	TD									
	INS									
	MT									
SR	SA									
	R									

		Intervals								
		1	2	3	4	5	6	7	8	9
TI	TD									
	INS									
	MT									
SR	SA									
	R									

		Intervals								
		1	2	3	4	5	6	7	8	9
TI	TD									
	INS									
	MT									
SR	SA									
	R									

Figure 4.2 SOSMART coding sheet

SOSMART Summary Scores

Category	Code	Number of category intervals	Total number of intervals for observation period	Percentage of occurrence	Frequency of events
Teacher Direct (TD)					
Classroom Teacher	CT				
Other	O				
None	N				
Subtotal					
Instruction (INS)					
Teacher-led	T				
Technology-led	C				
Subtotal					
Movement Type (MT)					
Reward/Incentive	R				
Opening Activity	O				
Teacher Directed Transition	TT				
Other Movement (non-academic)	OMna				
Other Movement (academic)	OMa				
Subtotal					
Students Active (SA)					
Whole class	W				
Part class	P				
Small group	G				
None	N				
Subtotal					
As a Result of What (R)					
Physical Environment	E				
Non-Teacher Directed Transition	NT				
Subtotal					
Grand total				100%	

Figure 4.3 SOSMART scoring summary

Note: Adapted from Observation Recording Record of Physical Educator's Teaching Behavior (ORRPETB), Stewart (1989) in van der Mars (1989)

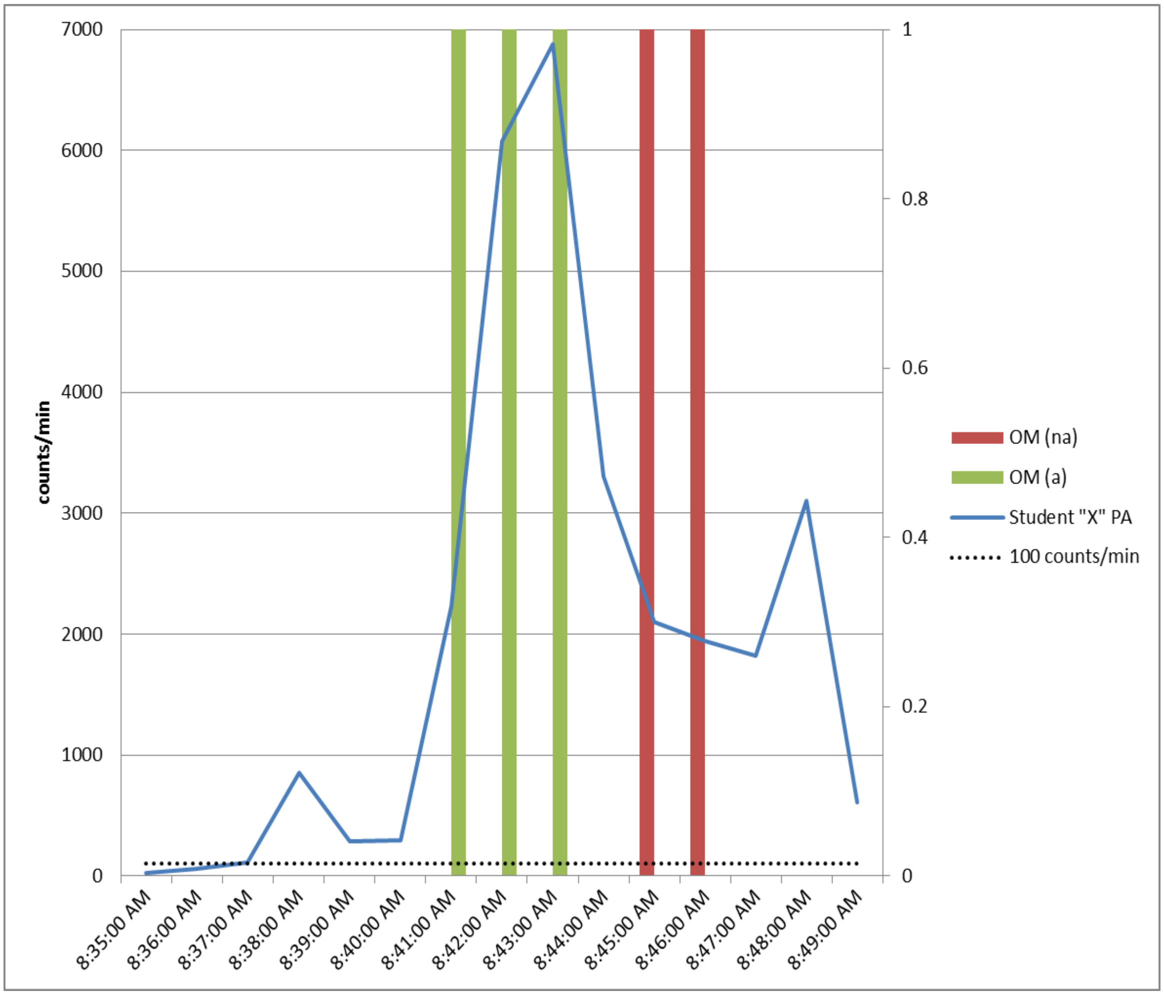

Figure 4.4 SOSMART construct validity of OM (na) and OM (a)

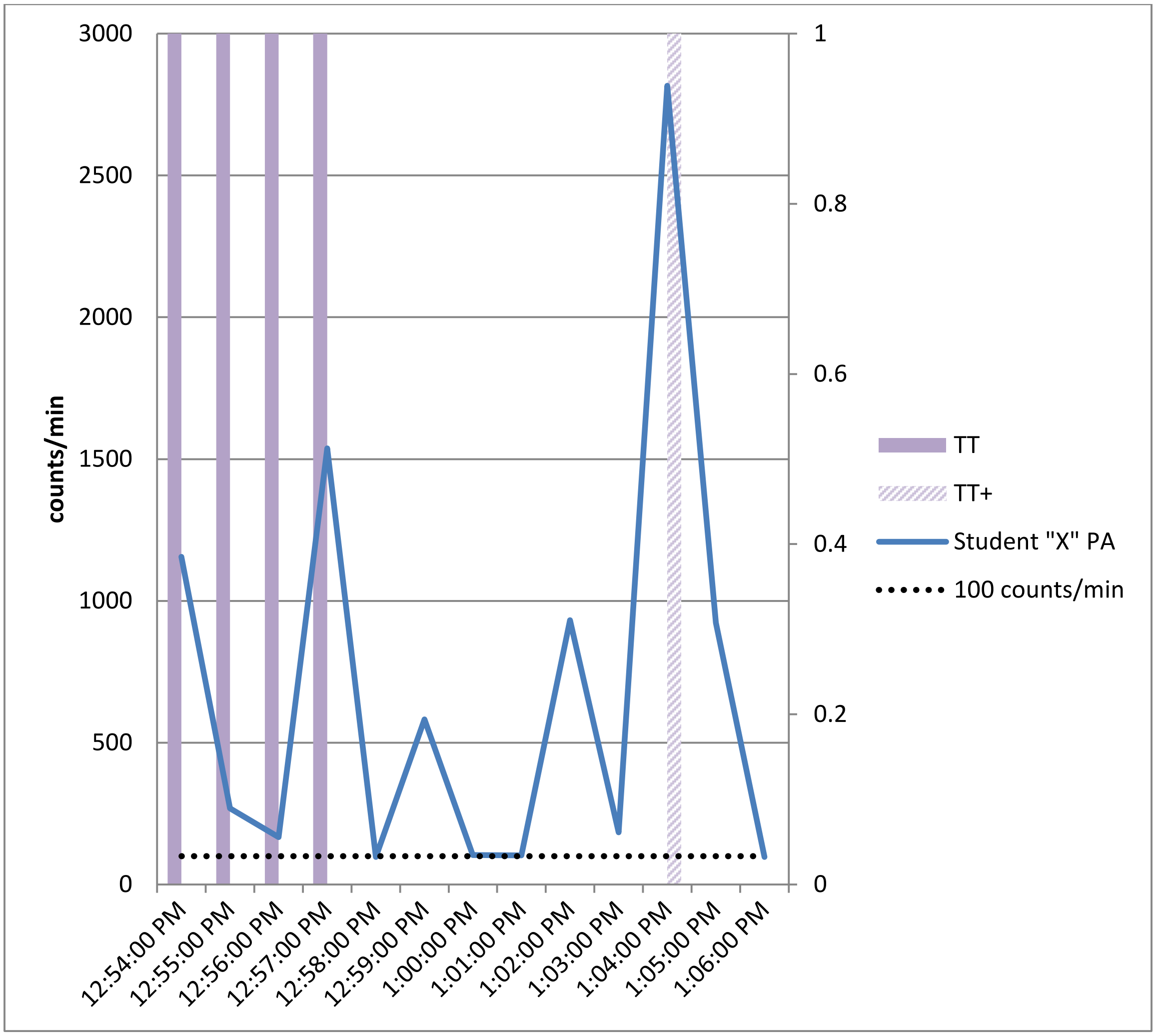

Figure 4.5 SOSMART construct validity of TT

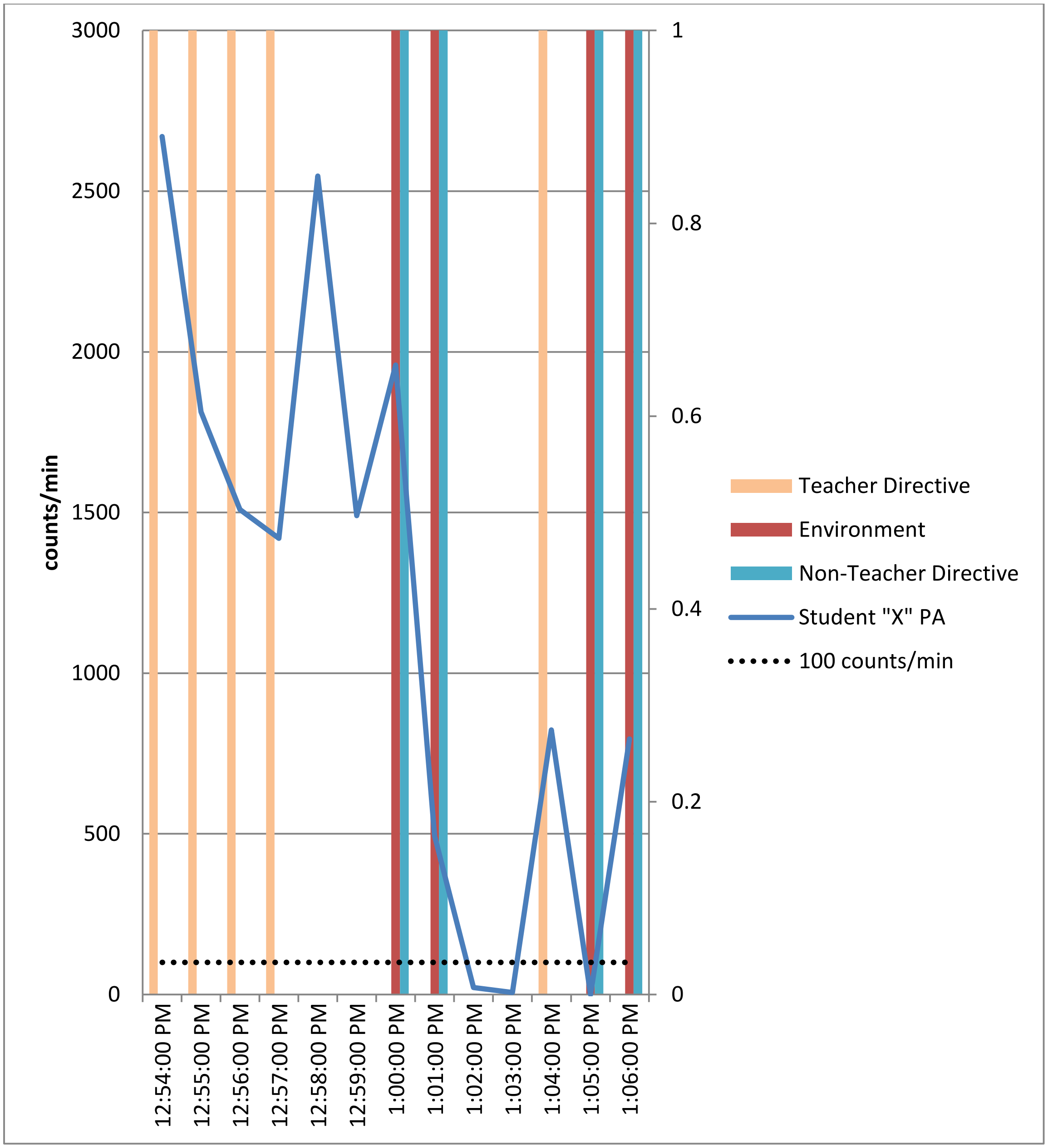

Figure 4.6 SOSMART construct validity of TD and NT

CHAPTER 5

Discussion

The contribution of this dissertation to advancing the knowledge base informing CSPAP adoption is two-fold. First, by conducting a systematic review and meta-analysis of multi-component school-based PA interventions, empirical evidence was generated to create a rationale for the continued pursuit of CSPAP effectiveness.

While the overall effect size was minimal, reasons for the lack of effectiveness are unclear (Russ, et al., 2015). Results from the first study indicate that pursuing CSPAPs is still a worthy endeavor but strategies within and across components need to be analyzed. Intervention components were not always in alignment with national recommendations (i.e. QPE was not taught by a certified professional), fidelity of implementation relied on self-reports, and staff trainings revealed inconsistencies. Targeting the quality of each intervention component, as well as measuring fidelity of implementation through developing component-specific objective measures, are strategies that could help enhance program effectiveness. Interventions that reflect all five components of the CSPAP model, align with current recommendations, provide detailed descriptions of intervention component design and implementation, and demonstrate dynamic collaboration across all five components are needed.

The second way this dissertation contributes to advancing the knowledge base for CSPAP efforts is by providing a component-specific (PADSD) objective measure of implementation. Within the school day, students spend a majority of their time with a classroom teacher across several different settings (i.e. classroom, lunch, recess), thereby insinuating the importance of CT involvement in PA promotion (in intervention and non-intervention contexts). Many multi-component school-based PA interventions have targeted the classroom as one of the settings to intervene (Russ, et al., 2015); however, the only measures of implementation fidelity reported in the classroom were self-reports. Through developing a systematic observation tool designed to capture the frequency and variety of strategies teachers use to integrate movement in the classroom setting (SOSMART), CSPAP efforts within this setting can now base policy and practice decisions on objective measurement data.

Data generated from utilizing SOSMART can be used to enhance pre-service teacher education, in-service teacher professional development, and future CSPAP research efforts. Teacher training (i.e. preservice and inservice CTs) can now utilize the MI terms presented in SOSMART as a common language to discuss MI strategies, and researchers can begin to explore which MI strategies may be more or less effective for CTs practicing in certain contexts. SOSMART can also be used to advance CSPAP research through providing descriptive data on the nature of MI in classrooms and objectively measuring implementation fidelity.

This dissertation represents one of the early efforts of CSPAP research. The combined impact of the studies herein results in a significant contribution to advancing the knowledge base needed for CSPAPs through providing empirical evidence and objective measures on which CSPAP efforts can now be grounded. Combined with other component-specific objective measures, continued use of SOSMART can contribute to the first efforts of national surveillance data documenting the implementation and effectiveness of CSPAPs. This, in turn, can facilitate the creation of a national benchmark for MI and/or reducing sedentarism in the academic classroom, which may result in a trickle-down effect influencing the criteria on which administrators evaluate CTs in the future. These contributions create a driving force behind CSPAP, moving forward the potential and possibility of wide scale program adoption.

REFERENCES

Ahamed, Y., Macdonald, H., Reed, K., Naylor, P. J., Liu-Ambrose, T., & Mckay, H. (2007). School-based physical activity does not compromise children's academic performance. *Medicine and Science in Sports and Exercise, 39*(2), 371.

American Alliance for Health, Physical Education, Recreation, and Dance [AAHPERD]. (2011). 2011 Comprehensive School Physical Activity Survey.

Bailey, D. P., Fairclough, S. J., Savory, L. A., Denton, S. J., Pang, D., Deane, C. S., & Kerr, C. J. (2012). Accelerometry-assessed sedentary behaviour and physical activity levels during the segmented school day in 10–14-year-old children: the HAPPY study. *European Journal of Pediatrics*, *171*(12), 1805-1813.

Baranowski, T., & Jago, R. (2005). Understanding the mechanisms of change in children's physical activity programs. *Exercise and Sport Sciences Reviews*, *33*(4), 163-168.

Bartholomew, J. B., & Jowers, E. M. (2011). Physically active academic lessons in elementary children. *Preventive Medicine*, *52*, S51-S54.

Bassett, D. R., Fitzhugh, E. C., Heath, G. W., Erwin, P. C., Frederick, G. M., Wolff, D. L., ... & Stout, A. B. (2013). Estimated energy expenditures for school-based policies and active living. *American Journal of Preventive Medicine*, *44*(2), 108-113.

Beighle, A., Erwin, H. E., Beets, M. W., Morgan, C. F., & Le Masurier, G. (2010). America on the move: School-based physical activity promotion. *International Journal of Physical Education*, *47*(2), 2-16.

Biddle, S. J., Gorely, T., & Stensel, D. J. (2004). Health-enhancing physical activity and sedentary behaviour in children and adolescents. *Journal of Sports Sciences*, *22*(8), 679-701.

Brown, W. H., Pfeiffer, K. A., McIver, K. L., Dowda, M., Almeida, J. M., & Pate, R. R. (2006). Assessing preschool children's physical activity: the Observational System for Recording Physical Activity in children-preschool version. *Research Quarterly for Exercise and Sport*, *77*(2), 167-176.

Caballero, B., Clay, T., Davis, S. M., Ethelbah, B., Rock, B. H., Lohman, T., . . . & Stephenson, L. (2003). Pathways: a school-based, randomized controlled trial for the prevention of obesity in American Indian school children. *The American Journal of Clinical Nutrition, 78*(5), 1030-1038.

Campbell, K., Waters, E., O'Meara, S., & Summerbell, C. (2001). Interventions for preventing obesity in childhood. A systematic review. *Obesity Reviews, 2*(3), 149-157.

Castelli, D. M., & Ward, K. (2012). Physical activity during the school day. *Journal of Physical Education, Recreation & Dance*, *83*(6), 20-29.

Centers for Disease Control and Prevention [CDC]. (2010). *The association between school based physical activity, including physical education, and academic performance.* Atlanta, GA: US Department of Health and Human Services.

Centers for Disease Control and Prevention [CDC]. (2013). Comprehensive School Physical Activity Programs: A guide for schools. Retrieved from http://www.cdc.gov/healthyyouth/physicalactivity/cspap.htm.

Cheffers, J. & Mancini, V. (1989). Cheffers' Adaptation of the Flanders' Interaction Analysis System (CAFIAS). In Darst, P., Zakrajsek, D., & Mancini, V. (eds.), *Analyzing Physical Education and Sport Instruction.* Champaign, IL: Human Kinetics.

Cothran, D. J., Kulinna, P. H., & Garn, A. C. (2010). Classroom teachers and physical activity integration. *Teaching and Teacher Education*, *26*(7), 1381-1388.

Cradock, A., Barrett, J., Carter, J., McHugh, A., Sproul, J., Russo, E., … & Gortmaker, S. (2014). Impact of the Boston Active School Day policy to promote physical activity among children. *American Journal of Health Promotion, 28*(3) S54-S64.

Damschroder, L. J., Aron, D. C., Keith, R. E., Kirsh, S. R., Alexander, J. A., & Lowery, J. C. (2009). Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. *Implementation Science*, *4*(1), 50.

DerSimonian, R., & Laird, N. M. (1986). Meta-analysis in Clinical Trials. *Controlled Clinical Trials, 7*, 177-178.

Dietz, W. H. (2001). The obesity epidemic in young children. *British Medical Journal, 322*(7282), 313-314.

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.

Donnelly, J. E., & Lambourne, K. (2011). Classroom-based physical activity, cognition, and academic achievement. *Preventive Medicine*, *52*, S36-S42.

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.

Dunn, L. L., Venturanza, J. A., Walsh, R. J., & Nonas, C. A. (2012). An observational evaluation of Move-To-Improve, a classroom-based physical activity program, New York City Schools, 2010. *Preventing Chronic Disease*, *9*.

Eccles, M. P., & Mittman, B. S. (2006). Welcome to implementation science. *Implementation Science*, *1*(1), 1-3.

Elliot, E., Erwin, H., Hall, T., & Heidorn, B. (2013) Comprehensive school physical activity programs: Helping all students achieve 60 minutes of physical activity each day. *Journal of Physical Education, Recreation & Dance, 84*(9), 9-15

Elmakis, G. S. (2010). *Survey of physical activity in elementary school classrooms in the state of Virginia.* (Doctoral dissertation). The College of William and Mary, Virginia.

Engbers, L. H., van Poppel, M. N., Chin A Paw, M. J., & van Mechelen, W. (2005). Worksite health promotion programs with environmental changes: a systematic review. *American Journal of Preventive Medicine*, *29*(1), 61-70.

Erwin, H. E. (2009). Implementing classroom-based physical activity. *Research Quarterly for Exercise and Sport*, 80(1), A10-A11.

Erwin, H. E., Beighle, A., Morgan, C. F., & Noland, M. (2011). Effect of a low-cost, teacher-directed classroom intervention on elementary students' physical activity. *Journal of School Health*, *81*(8), 455-461.

Erwin, H., Beighle, A., Carson, R. L., & Castelli, D. M. (2013). Comprehensive school-based physical activity promotion: A review. *Quest*, *65*(4), 412-428.

Evenson, K. R., Ballard, K., Lee, G., & Ammerman, A. (2009). Implementation of a school-based state policy to increase physical activity. *Journal of School Health, 79*(5), 231-238.

Flanders, N. A. (1970). *Analyzing teacher behavior*. Addison-Wesley P. C.

Flanders, N. A. (1976). Interaction analysis and clinical supervision. *Journal of Research & Development in Education*.

Flodmark, C., Marcus, C., & Britton, M. (2006). Interventions to prevent obesity in children and adolescents: a systematic literature review. *International Journal of Obesity, 30*(4), 579-589.

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. *International Journal of Behavioral Nutrition and Physical Activity*, *5*(1), 36.

Goh, T.L., Hannon, J.C., Newton, M., Webster, C., Podlog, L., & Pillow, W. (2013). “I'll Squeeze It In”: Transforming preservice classroom teachers' perceptions toward movement integration in schools. *Action in Teacher Education, 35*(4), 286-300.

Gortmaker, S. L., Peterson, K., Wiecha, J., Sobol, A. M., Dixit, S., Fox, M. K., & Laird, N. (1999). Reducing obesity via a school-based interdisciplinary intervention among youth: Planet Health. *Archives of Pediatrics & Adolescent Medicine, 153*(4), 409-418.

Grieco, L. A., Jowers, E., & Bartholomew, J. B. (2009). Physically active academic lessons and time on task: the moderating effect of body mass index. *Medicine and Science in Sports and Exercise, 41*(10), 1921-6.

Guo, G., & Zhao, H. (2000). Multilevel modeling for binary data. *Annual review of sociology*, 441-462.

Hedges, L. V. (1982). Estimation of effect size from a series of independent experiments. *Psychological Bulletin, 92*, 490-499.

Higgins, J. P., Green, S., & Collaboration, C. (2008). *Cochrane Handbook for Systematic Reviews of Interventions* (Vol. 5): Wiley Online Library.

Hoelscher, D. M., Springer, A. E., Ranjit, N., Perry, C. L., Evans, A. E., Stigler, M., & Kelder, S. H. (2010). Reductions in child obesity among disadvantaged school children with community involvement: the Travis County CATCH Trial. *Obesity, 18,* S36-S44.

Holt, E., Bartee, T., & Heelan, K. (2013). Evaluation of a policy to integrate physical activity into the school day. *Journal of Physical Activity & Health, 10*(4).

Howie, E. K. (2013). *Classroom exercise breaks and educational outcomes in elementary school students.* (Doctoral dissertation). University of South Carolina, South Carolina.

Howie, E. K., Newman-Norlund, R. D., & Pate, R. R. (2014). Smiles count but minutes matter: responses to classroom exercise breaks. *American Journal of Health Behavior*, *38*(5), 681-689.

Huedo-Medina, T. B., Sanchez-Meca, J., Marin-Martinez, F., & Botella, J. (2006). Assessing heterogeneity in meta-analysis: Q statistic or I2 index? *Psychological Methods, 11*(2), 193-206.

Institute of Medicine [IOM]. (2013). *Educating the student body: Taking physical activity and physical education to school*. Washington, DC: The National Academies Press.

Janulewicz, P. (2008). *Replacement of chairs with fitness balls in elementary school classrooms to increase daily physical activity.* (Master's thesis). University of Nebraska, Kearney.

Katz, D. L., Cushman, D., Reynolds, J., Njike, V., Treu, J. A., Katz, C., …, & Smith, E. (2010). Putting physical activity where it fits in the school day: Preliminary results of the ABC (Activity Bursts in the Classroom) for fitness program. *Preventing Chronic Disease*, *7*(4).

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*, S43-S50.

Kohl, H. W., Moore, B. M., Sutton, A. W., Kibbe, D. L., & Schneider, D. C. (2001). A curriculum-integrated classroom physical activity promotion tool for elementary schools: Teacher Evaluation of Take 10!™. *Medicine & Science in Sports & Exercise*, *33*(5), S179.

Lee, S. M., Burgeson, C. R., Fulton, J. E., & Spain, C. G. (2007). Physical education and physical activity: results from the School Health Policies and Programs Study 2006. *Journal of School Health, 77*(8), 435-463.

Liberati, A., Altman, D. G., Tetzlaff, J., Mulrow, C., Gøtzsche, P. C., Ioannidis, J. P., . . . & Moher, D. (2009). The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *Annals of Internal Medicine, 151*(4), W-65-W-94.

Lonsdale, C., Rosenkranz, R. R., Peralta, L. R., Bennie, A., Fahey, P., & Lubans, D. R. (2013). A systematic review and meta-analysis of interventions designed to increase moderate-to-vigorous physical activity in school physical education lessons. *Preventive Medicine*, *56*(2), 152-161.

Luepker, R. V., Perry, C. L., McKinlay, S. M., Nader, P. R., Parcel, G. S., Stone, E. J., . . . & Johnson, C. C. (1996). Outcomes of a field trial to improve children's dietary patterns and physical activity. *Journal of the American Medical Association, 275*(10), 768-776.

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.

Mahar, M. T. (2011). Impact of short bouts of physical activity on attention-to-task in elementary school children. *Preventive Medicine*, *52,* S60-S64.

Matthews, C. E., Chen, K. Y., Freedson, P. S., Buchowski, M. S., Beech, B. M., Pate, R. R., & Troiano, R. P. (2008). Amount of time spent in sedentary behaviors in the United States, 2003–2004. *American Journal of Epidemiology*, *167*(7), 875-881.

McGraw, S. A., Sellers, D., Stone, E., Resnicow, K. A., Kuester, S., Fridinger, F., & Wechsler, H. (2000). Measuring implementation of school programs and policies to promote healthy eating and physical activity among youth. *Preventive Medicine*, *31*(2), S86-S97.

McKenzie, T. L., & Kahan, D. (2008). Physical activity, public health, and elementary schools. *The Elementary School Journal*, *108*(3), 171-180.

McKenzie, T. L., Cohen, D. A., Sehgal, A., Williamson, S., & Golinelli, D. (2006). System for Observing Play and Recreation in Communities (SOPARC): reliability and feasibility measures. *Journal of Physical Activity & Health*, *3*, S208.

McKenzie, T. L., Marshall, S. J., Sallis, J. F., & Conway, T. L. (2000). Leisure-time physical activity in school environments: an observational study using SOPLAY. *Preventive Medicine*, *30*(1), 70-77.

McKenzie, T. L., Sallis, J. F., & Nader, P. R. (1992). SOFIT: System for observing fitness instruction time. *Journal of Teaching in Physical Education*, *11*(2).

McKenzie, T. L. & van der Mars, H. (2015). Top 10 Research Questions Related to Assessing Physical Activity and Its Contexts Using Systematic Observation. *Research Quarterly for Exercise and Sport, 86,* 13-29.

Metcalf, B., Henley, W., & Wilkin, T. (2012). Effectiveness of intervention on physical activity of children: systematic review and meta-analysis of controlled trials with objectively measured outcomes (EarlyBird 54). *British Medical Journal*, *345*.

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Annals of Internal Medicine*, *151*(4), 264-269.

Morris, S. B., & DeShon, R. P. (2002). Combining effect size estimates in meta-analysis with repeated measures and independent-groups designs. *Psychological Methods, 7*(1), 105-125.

National Association for Sport and Physical Education [NASPE] (2013). *Comprehensive school physical activity programs: Helping all students achieve 60 minutes of physical activity each day* [Position statement]. Reston, VA: Author. Retrieved from http://www.shapeamerica.org/advocacy/positionstatements/pa/loader.cfm?csModule=security/getfile&pageid=4726

National Physical Activity Plan Alliance. (2012). National Physical Activity Plan.

Naylor, P. J., Macdonald, H. M., Zebedee, J. A., Reed, K. E., & McKay, H. A. (2006). Lessons learned from Action Schools! BC—an 'active school'model to promote physical activity in elementary schools. *Journal of Science and Medicine in Sport*, *9*(5), 413-423.

Neumark-Sztainer, D., Story, M., Hannan, P. J., & Rex, J. (2003). New Moves: a school-based obesity prevention program for adolescent girls. *Preventive Medicine, 37*(1), 41-51.

Neumark-Sztainer, D. R., Friend, S. E., Flattum, C. F., Hannan, P. J., Story, M. T., Bauer, K. W., . . . & Petrich, C. A. (2010). New Moves—preventing weight-related problems in adolescent girls: a group-randomized study. *American Journal of Preventive Medicine, 39*(5), 421-432.

Orlowski, M. A., & Hart, A. (2010). Go! Including Movement during Routines and Transitions. *Young Children, 65*(5), 88-93.

Orlowski, M., Lorson, K., Lyon, A., & Minoughan, S. (2013). My classroom physical activity pyramid: A tool for integrating movement into the classroom. *Journal of Physical Education, Recreation, & Dance, 84*(9), 47-51.

Pangrazi, R. P., Beighle, A., Vehige, T., & Vack, C. (2003). Impact of Promoting Lifestyle Activity for Youth (PLAY) on children's physical activity. *Journal of School Health, 73*(8), 317-321.

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

Pate, R. R., Ward, D. S., Saunders, R. P., Felton, G., Dishman, R. K., & Dowda, M. (2005). Promotion of physical activity among high-school girls: a randomized controlled trial. *Journal Information, 95*(9).

Pope, R. P., Coleman, K. J., Gonzalez, E. C., Barron, F., & Heath, E. M. (2002). Validity of a revised system for observing fitness instruction time (SOFIT). *Pediatric Exercise Science*, *14*(2), 135-146.

Prince, S. A., Adamo, K. B., Hamel, M. E., Hardt, J., Gorber, S. C., & Tremblay, M. (2008). A comparison of direct versus self-report measures for assessing physical activity in adults: a systematic review. *International Journal of Behavioral Nutrition and Physical Activity*, *5*(1), 56.

Proctor, E. K., Landsverk, J., Aarons, G., Chambers, D., Glisson, C., & Mittman, B. (2009). Implementation research in mental health services: An emerging science with conceptual, methodological, and training challenges. *Administration and Policy in Mental Health and Mental Health Services Research*, *36*(1), 24-34.

Ridgers, N. D., Stratton, G., & Fairclough, S. J. (2006). Physical activity levels of children during school playtime. *Sports Medicine*, *36*(4), 359-371.

Robinson, T. N. (1999). Reducing children's television viewing to prevent obesity: a randomized controlled trial. *Journal of the American Medical Association*, *282*(16), 1561-1567.

Russ, L., Webster, C.A., Beets, M.W., & Phillips, D.S. (2015). Systematic review and meta-analysis of multicomponent interventions through schools to increase physical activity. *Journal of Physical Activity and Health.*

Sallis, J. F., McKenzie, T. L., Alcaraz, J. E., Kolody, B., Faucette, N., & Hovell, M. F. (1997). The effects of a 2-year physical education program (SPARK) on physical activity and fitness in elementary school students. Sports, Play and Active Recreation for Kids. *American Journal of Public Health, 87*(8), 1328-1334.

Sallis, J. F., McKenzie, T. L., Conway, T. L., Elder, J. P., Prochaska, J. J., Brown, M., . . . & Alcaraz, J. E. (2003). Environmental interventions for eating and physical activity: a randomized controlled trial in middle schools. *American Journal of Preventive Medicine, 24*(3), 209-217.

Salmon, J. O., Ball, K., Crawford, D., Booth, M., Telford, A., Hume, C., ... & Worsley, A. (2005). Reducing sedentary behaviour and increasing physical activity among 10-year-old children: overview and process evaluation of the 'Switch-Play'intervention. *Health Promotion International*, *20*(1), 7-17.

Salmon, J. (2010). Novel strategies to promote children's physical activities and reduce sedentary behavior. *Journal of Physical Activity & Health*, *7*(3), S299.

Seo, D.-C., King, M., Kim, N., Sovinski, D., Meade, R., Lederer, A. (2013). Predictors for moderate- and vigorous-intensity physical activity during an 18-month coordinated school health intervention. *Preventive Medicine*. doi: http://dx.doi.org/10.1016.j.ypmed.2013.06.024

Skrade, M. (2013). *Integrated classroom physical activity: Examining perceived need satisfaction and academic performance in children*. (Doctoral dissertation). Iowa State University, Iowa.

Skrade, M. & Vazou, S. (2013). Integrated classroom physical activity: Examining perceived need satisfaction and academic performance in children. *Journal of Sport & Exercise Psychology,* 35, S8-S12.

South Carolina Department of Education (2013). *E-rate Free and Reduced Meal Eligibility Data.* Retrieved from https://ed.sc.gov/data/erate/index.cfm

Springer, A. E., Kelder, S. H., Ranjit, N., Hochberg-Garrett, H., Crow, S., & Delk, J. (2012). Promoting physical activity and fruit and vegetable consumption through a community-school partnership: the effects of Marathon Kids® on low-income elementary school children in Texas. *Journal of Physical Activity & Health, 9*(5), 739-753.

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.

Story, M., Hannan, P. J., Fulkerson, J. A., Rock, B. H., Smyth, M., Arcan, C., & Himes, J. H. (2012). Bright Start: Description and Main Outcomes From a Group-Randomized Obesity Prevention Trial in American Indian Children. *Obesity, 20*(11), 2241-2249.

Sturm, R. (2005). Childhood obesity—what we can learn from existing data on societal trends, part 1. *Preventing Chronic Disease 2*(2), 1–9.

Till, J., & Ferkins, L. (2014). First-hand participation: illuminating teachers' self-perceptions of physical activity delivery. *Asia-Pacific Journal of Health, Sport and Physical Education, 5*(1), 39-54.

United States Department of Health and Human Services [USDHHS]. (2008). *Physical Activity Guidelines for Americans.* Retrieved from http://www.health.gov/paguidelines

United States Department of Health and Human Services [USDHHS]. (2012). *Physical Activity Guidelines for Americans Mid-Course Report: Strategies to Increase Physical Activity Among Youth*. Retrieved from http://www.health.gov/paguidelines/midcourse/pag-mid-course-report-final.pdf.

van der Mars, H. (1989). Systematic observation: An introduction. In Darst, P., Zakrajsek, D., & Mancini, V. (eds.), *Analyzing Physical Education and Sport Instruction.* Champaign, IL: Human Kinetics.

van der Mars, H. (1989b). Basic recording tactics. In Darst, P., Zakrajsek, D., & Mancini, V. (eds.), *Analyzing Physical Education and Sport Instruction.* Champaign, IL: Human Kinetics.

van Sluijs, E. M., van Poppel, M. N., & van Mechelen, W. (2004). Stage-based lifestyle interventions in primary care: are they effective? *American Journal of Preventive Medicine, 26*(4), 330-343.

van Sluijs, E., McMinn, A., & Griffin, S. (2007). Effectiveness of interventions to promote physical activity in children and adolescents: systematic review of controlled trials. *British Medical Journal, 335*(7622), 677-678.

Weaver, R. G., Beets, M. W., Webster, C., & Huberty, J. L. (2014). System for observing staff promotion of activity and nutrition (SOSPAN). *Journal of Physical Activity & Health, 11*(1).

Webber, L. S., Catellier, D. J., Lytle, L. A., Murray, D. M., Pratt, C. A., Young, D. R., . . . & Jobe, J. B. (2008). Promoting physical activity in middle school girls: Trial of

Activity for Adolescent Girls. *American Journal of Preventive Medicine, 34*(3), 173-184.

Webster, C. A., Caputi, P., Perreault, M., Doan, R., Weaver, G., & Doutis, P. (2013a). Elementary classroom teachers' adoption of physical activity promotion in the context of a statewide policy: An innovation diffusion and socio-ecologic perspective. *Journal of Teaching in Physical Education*, *32*(4), 419-440.

Webster, C. A., Wellborn, B., Hunt, K., LaFleche, M., Cribbs, J., & Lineberger, B. (2013b). MPOWER: An observation system for assessing coach autonomy support in high school varsity boys' soccer practices. *International Journal of Sports Science and Coaching*, *8*(4), 741-754.

Webster, C., Russ, L., Vazou, S., Goh, T., & Erwin, H. (2015). Integrating Movement in Academic Classrooms: Understanding, Applying, and Advancing the Knowledge Base. *Obesity Reviews*

Williamson, D. A., Copeland, A. L., Anton, S. D., Champagne, C., Han, H., Lewis, L., ... & Ryan, D. (2007). Wise Mind Project: A school-based environmental approach for preventing weight gain in children. *Obesity*, *15*(4), 906-917.

Woods, C. D. (2011). *Evaluation of Instant Recess® exercise breaks as a means for implementing LAUSD physical activity policy in elementary schools.* (Doctoral dissertation). University of California, Los Angeles.

Young, D. R., Phillips, J. A., Yu, T., & Haythornthwaite, J. A. (2006). Effects of a life skills intervention for increasing physical activity in adolescent girls. *Archives of Pediatrics & Adolescent Medicine*, *160*(12), 1255.

Appendix A: Permission to Reprint

From: Doug Hoepker <dough@hkusa.com>
Sent: Monday, June 08, 2015 12:35 PM
To: RUSS, LAURA
Subject: RE: Your JPAH article is now In Press

Laura,

You have our permission to deposit the article, accepted to JPAH, with your university in order to complete your dissertation.

Doug Hoepker
Human Kinetics
Senior Journals Managing Editor
JPAH, *RSJ*, & *JLAS*
Follow JPAH on twitter: @JPAHjournal
Follow JLAS on twitter: @JLASjournal
p: 217-351-5076, x2404
f: 217-239-1215
dough@hkusa.com