Summer 2021

Three Studies Investigating Comprehensive School Physical Activity Program-Aligned Opportunities To Enhance Students' Physical Education Learning

Jongho Moon

THREE STUDIES INVESTIGATING COMPREHENSIVE SCHOOL PHYSICAL ACTIVITY PROGRAM-ALIGNED OPPORTUNITIES TO ENHANCE STUDENTS' PHYSICAL EDUCATION LEARNING

by

Jongho Moon

Bachelor of Arts
Gwangju National University of Education, 2012

Master of Arts
Gwangju National University of Education, 2017

Submitted in Partial Fulfillment of the Requirements

For the Degree of Doctor of Philosophy in

Physical Education

College of Education

University of South Carolina

2021

Accepted by:

Collin A. Webster, Major Professor

Ali Brian, Committee Member

David F. Stodden, Committee Member

Kelly Lynn Mulvey, Committee Member

Tracey L. Weldon, Interim Vice Provost and Dean of the Graduate School

DEDICATION

This dissertation is dedicated to my wife, Yoonjoo Lee, who has been a constant source of love, support, and encouragement during the challenges of graduate school and life. I am truly thankful to have her in my life. This work is also dedicated to my daughter, Seoyeon Moon, who continues to be the light of my life. Without their unconditional love and sacrifice, this dissertation would not have been made possible. Thank you for everything.

I also dedicate this dissertation to my beloved parents, Jeongjin Moon and Insuk Choi, who have always supported me unconditionally and who have inspired me to strive for strength, diligence, and patience. Their support ultimately taught me to keep moving forward and to continue pursuing all I aspire to achieve.

ACKNOWLEDGEMENTS

I would like to express my gratitude to my supervisor Dr. Collin A. Webster for his countless hours of reflecting, reading, writing, and encouraging throughout the entire process. You have been my mentor, my teacher, and my friend; you always believe in me, which has made me never give up on this tough journey. This will be one of the best memories of my life. Thank you for running with me to the finish line—I hope our paths continue to intersect in the future.

My doctoral committee members, Drs. Ali Brian, David Stodden, and Kelly Lynn Mulvey—thank you for your expertise, guidance, and time, all which strengthened the quality of not only this dissertation but myself as a researcher.

To all my graduate assistant friends, Matthew Patey, Jenna Fisher, Adam Pennell, Emily Gilbert, Alex Stribing, Hayes Bennett, Stephen Griffin, and Lori Flick—thank you for being a part of my dynamic adventure, for the willingness to share priceless experiences, and for always helping me see the light at the end of the tunnel.

A special thanks to Karie Orendorff—I would like to sincerely express my gratitude for your invaluable support. I also would like to give special thanks to my friend, Christina Xan, for many hours of proofreading. Finally, thanks to my personal mentor, Dr. Dongwoo Lee, and all my FC GNUE friends, you have always uplifted my hopes, dreams, and journeys, even from the other side of the world.

ABSTRACT

Schools have been identified as key venues that have a positive impact on students' development through physical activity (PA) and physical education (PE) programming. A comprehensive school physical activity program (CSPAP) is the Centers for Disease Control and Prevention's [CDC] national framework for school-based PA (CDC, 2019). A CSPAP was conceptualized as a strategic approach to leverage the full range of resources needed to meet two broad goals: (a) provide school-aged youth with the opportunity to meet the nationally recommended 60 minutes of daily PA, and (b) achieve Society of Health and Physical Educators (SHAPE) America's National Standards for K-12 learners (SHAPE America, 2014). While it is well established that a CSPAP can facilitate children's PA opportunities, little is known about the extent to which the different components of a CSPAP can support students' achievement of the PE standards.

Originally, the purpose of this dissertation was to investigate the potential of opportunities outside of traditional PE to bolster students' PE learning. The dissertation was based on the idea that a CSPAP can be used to enhance and accelerate what students learn in PE, including physical, cognitive, and social and emotional skills. From this perspective, all CSPAP components are connected not only through PA experiences but also PE learning. Three studies were proposed for this dissertation, including (a) a systematic review and meta-analysis of the effectiveness of CSPAP-aligned PA interventions in increasing elementary children's motor competence (MC), in line with

SHAPE America's (2014) Standard 1 (Study 1); (b) a systematic review and meta-analysis of the effectiveness of CSPAP-aligned PA interventions on the development of elementary children's social and emotional learning (SEL), in line with SHAPE America's Standard 4 and 5 (Study 2); and (c) a school-based pilot intervention examining the efficacy of classroom movement integration (MI) to support elementary children's MC and SEL.

Due to the onset of the COVID-19 pandemic during this dissertation, conducting the intervention for Study 3 was not possible because of school closures and the author and his committee agreed to substitute the intervention with the development of a new research measure for assessing online physical education (OLPE), which was considered appropriate in the context of the pandemic. As such, the revised purpose of this dissertation was to examine CSPAP-aligned strategies, particularly beyond traditional PE, for supporting students' PE learning. While no changes were made to Studies 1 or 2, the updated purpose of Study 3 was to develop a systematic observation instrument – the System for Observing Virtual Real Time Lessons in Physical Education (SOVRTL-PE) – for assessing synchronous OLPE lessons in K-12 PE.

Studies 1 and 2 followed the updated Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Page et al., 2021) and synthesized the evidence of the effectiveness of PA interventions in increasing MC and SEL of children aged 5-12. The CSPAP framework was used to categorize the different intervention approaches. The results of both studies demonstrate that the PA intervention programs had a positive effect on the development of MC and SEL. In light of our results, extensive CSPAP-aligned PA intervention approaches appear to be promising

avenues for enhancing children's MC and SEL. Study 3 used five phases to develop and establish the reliability of the SOVRTL-PE. Specifically, the instrument was content validated, using an extensive literature search, observations of synchronous OLPE lessons, and consensus from a Delphi survey with several experts within the field. The result of Study 3 provides an overview of key features of synchronous OLPE and underscores the importance of optimizing family and community resources to better understand how to support students' learning during virtual PE instruction. In essence, this study provides one possibility for systematically assessing and improving professional practices amid increased shifts toward virtual teaching and learning alternatives in PE within the CSPAP framework.

Ultimately, understanding the potential of CSPAPs, whether facilitated in-person or remotely, for supporting students' PE learning fundamentally changes the way both researchers and practitioners might approach implementing such programs.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

CSPAP .... Comprehensive School Physical Activity Program

FCE .... Family and Community Engagement

FMS .... Fundamental Movement Skills

MC .... Motor Competence

MI .... Movement Integration

OLPE .... Online Physical Education

PA .... Physical Activity

PABAS .... Physical Activity Before and After School

PADS .... Physical Activity During School

PE .... Physical Education

RDM .... Responsible Decision Making

RS .... Relationship Skills

SA .... Self Awareness

SEL .... Social and Emotional Learning

SHAPE .... Society of Health and Physical Educators

SM .... Self Management

SoA .... Social Awareness

SOVRTL-PE .... System for Observing Virtual Real Time Lessons in Physical Education

CHAPTER 1

INTRODUCTION

Regular participation in physical activity (PA) is beneficial for the physical and mental health of school-aged youth (Centers for Disease Control and Prevention [CDC], 2010; Janssen & LeBlanc, 2010; Ströhle, 2009). Several studies have shown that PA is positively associated with improved physical fitness and brain function during development (Hillman et al., 2014; Kamijo et al., 2011). Engaging in PA also supports children and adolescents' academic (Bartholomew & Jowers, 2011; Donnelly & Lambourne, 2011; Riley et al., 2015) and social-emotional learning (SEL; Ang & Penney, 2013). Therefore, many school-based PA intervention programs have been purposefully developed with the aim to enhance PA levels of children and adolescents. Unfortunately, these programs have mostly been unsuccessful in increasing the PA of children and adolescents (Love et al., 2020) and only 24% of youth 5-17 years of age in the United States are meeting the nationally recommended minimum of 60 minutes of PA per day (National Physical Activity Plan Alliance, 2018).

Importance of Motor Competence in Physical Activity Participation

A key factor in children's PA participation is their motor competence (MC), specifically competence in fundamental motor skills (FMS; Barnett et al., 2009; Fisher et al., 2005; Logan et al., 2015; Stodden et al., 2008; Robinson et al., 2015). MC can be defined as the degree of skilled performance in a wide range of motor tasks as well as the movement quality, coordination and control underlying a particular motor outcome

(Burton & Miller, 1998; Gabbard, 2015). Additionally, FMS consists of three main constructs; locomotor (e.g., run, hop, jump, slide, gallop, leap); object control (e.g., strike, dribble, kick, throw, underarm roll, catch); and balance/stability skills (e.g., non-locomotor skills such as body rolling and twisting; Gallahue, Ozmun & Goodway, 2012). Higher gross MC attenuates the decline in PA throughout childhood (Lopes et al., 2011) and is important to physical development and PA across the developmental lifespan (Stodden et al., 2012). Several studies report that FMS competence is positively related to PA participation in childhood and youth (Haubenstricker & Seefeldt, 1986; Wrotniak et al. 2006; Lubans et al. 2010). In other words, if children do not learn to be competent in FMS, they may be less likely to engage in PA throughout their lives (Robinson et al., 2015). A number of recent studies add further support to this belief, highlighting positive associations between FMS competence and PA (Barnett et al., 2009; Barnett et al., 2016; Chan, Ha, & Ng., 2016).

Role of Motor Competence in Supporting Social and Emotional Learning: An Executive Functions Perspective

MC not only supports children's PA engagement but also may have unique benefits for SEL beyond the effects of PA (Biddle, Ciaccioni, Thomas, & Vergeer, 2019; Pesce et al., 2013; Vazou et al., 2019). SEL has become an increasingly popular topic in education research as today's children are placed under many expectations, such as academic achievement, cooperating with others, and resisting negative tendencies (Elias et al., 1997, 2002; Payton et al., 2000). Durlak et al. (2011), in a review of 34 articles, present that well-designed and effective SEL programs are associated with positive

social, emotional, behavioral, and academic outcomes for children. Hence, SEL should be purposefully taught to maximize student learning (Greenberg et al., 2003).

Emerging evidence suggests that context-specific PA (e.g., MC) in school settings enhances executive functioning (Best, 2010; Livesey et al., 2006; Mulvey et al., 2018; Pesce et al., 2016; Roebers & Kauer, 2009; Stökle & Hughes, 2016; van der Fels et al., 2015; van der Niet et al., 2016; Zysset et al., 2018), which in turn may be linked to developing SEL (Blair, 2006; Diamond & Lee, 2011; Diamond & Ling, 2015; Riggs et al., 2006). EFs are control processes which regulate behavior and cognition (Miyake & Friedman, 2012). They enable goal-directed behavior; thus, they are particularly needed in situations that involve planning, decision-making, and inhibition of inappropriate behavior (Banich, 2009; Best, 2010; Diamond, 2013). Several studies examining the functional structure of EFs have identified three main components: inhibition, working memory, and cognitive flexibility (Norman & Shallice, 1986; Miyake et al., 2000).

Studies show that PA and complex motor task engagement/movement stimulate specific areas of the brain (i.e., the dorsolateral prefrontal cortex, the contralateral cerebellum) as well as neural circuitry to help develop EFs (Best, 2010; Carey et al., 2005; Diamond, 2000; Jones et al., 1998; Serrien et al., 2007). This view is closely linked to the emerging line of chronic exercise and cognition research focused on pathways through which qualitative, rather than quantitative, PA characteristics can benefit cognitive functioning (Best, 2010; Myer et al., 2015; Pesce & Ben-Soussan, 2016; Pesce et al., 2021). A review by Best (2010) focused on studies investigating the relationship between EF and PA. These findings show that strategic play during group games, acute exercise, and chronic exercise have positive effects on cognitive patterns of those

engaged in EF tasks (Best, 2010). Cognitively engaging in PA (e.g., group activities, coordinative exercise) might be more beneficial for the improvement of EFs than simpler exercise (Best, 2010; Pesce, 2012; Schmidt, Jaeger, Egger, Roebers, & Conzelmann, 2015). For example, Schmidt et al. (2016) reported that the condition of an intervention that focused exclusively on aerobic exercise did not accrue the same cognitive benefits compared to the cognitively engaging PA conditions such as PA with high cognitive demands, sedentary with high cognitive demands, PA with low cognitive demands, and sedentary with low cognitive demands. In contrast, Mulvey et al. (2018) presented the effectiveness of the Successful Kinesthetic Instruction for Preschoolers (SKIP) motor development intervention program in improving children's executive function. SKIP involves a developmental array of gross motor tasks divided into locomotor (e.g., run, jump, leap, hop, gallop, and slide) and object control (e.g., throw, catch, kick, dribble, strike, and roll) skills. All tasks feature within-task variations and all participants receive their own equipment. Both support all learners' individual needs. However, purposefully designed PA is still rarely explored in intervention programs that attempt to promote social-emotional competence (Riggs et al., 2006).

School Programming to Develop Motor Competence

As Robinson and Goodway (2009) state, "FMS must be learned, practiced, and reinforced." (p. 533). That is, children need to learn and practice to gain competency (Brian et al., 2017; Clark, 1994; Logan, Robinson, Wilson, & Lucas, 2012). Therefore, the teaching of MC programs is important, as MC does not develop naturally (Brian et al., 2017; Hardy et al., 2010; Logan et al., 2012). The development of children and adolescents' MC is a primary goal of school PE and is considered foundational to

promoting lifetime participation in PA (Society of Health and Physical Educators [SHAPE] America, 2014). Standard 1 of the national standards for K-12 PE (SHAPE America, 2014) states, "The physically literate individual demonstrates competency in a variety of motor skills and movement patterns." However, PE programs face numerous limitations (e.g., limited allocated curriculum time, lack of well-defined policies, disparities between policy requirements and actual implementation; Hardman, 2011), and less than half of students appear to be achieving Standard 1 (Hastie, 2017). For instance, only 15% of elementary schools require or recommend that students take PE at least three times per week for the entire school year, and more than one third (35.2%) of school districts do not have a policy requiring elementary schools to provide regularly scheduled recess (National Physical Activity Plan Alliance, 2018).

A possible mechanism for extending students' motor skill learning beyond PE is a comprehensive school physical activity program (CSPAP; CDC, 2013; NASPE, 2008; SHAPE America, 2015). The CSPAP model includes five components: (a) quality of PE, (b) PA during school, (c) PA before and after school, (d) staff involvement, and (e) family and community engagement (CDC, 2013, 2015; NASPE, 2008; SHAPE America, 2015). The CSPAP model has become a widely disseminated paradigm in the literature concerning whole-of-school approaches to youth PA promotion (Webster et al., 2020). Through such an approach, additional PA opportunities before, during, and/or after school are provided to (a) help students meet the nationally recommended 60 or more minutes of daily PA, and (b) support students' learning of the knowledge and skills needed to participate in a lifetime of PA. While researchers have established that school-related PA opportunities beyond PE can increase students' participation in PA (Erwin et

al., 2013), less is known about the extent to which such opportunities may directly support MC or indirectly support SEL.

Untapped Potential of the General Education Classroom Setting

Movement Integration (MI), defined as incorporating PA at any level of intensity into daily classroom settings, has gained traction in research and national recommendations as a key strategy for providing children with supplemental PA during school (Russ et al., 2015; Webster et al., 2015), and may also serve as a potential strategy to support students' MC (Webster et al., 2020). The targeted context for MI is most often elementary general education classroom contexts, where students receive instruction from a classroom teacher in a range of academic subjects, such as math, science, social studies, etc. Examples of MI include teacher-directed physical activity breaks between/during academic lessons (e.g., showing a GoNoodle video or leading a brief exercise routine), learning experiences that involve students being physically active (e.g., having students act out stories or move into groups to demonstrate answers to division problems), and routine transitions that require students to move from one part of the classroom to another (e.g., having students skip from desks to carpet) (Moon & Webster, 2019; Russ et al., 2015; Webster et al. 2015).

MI has garnered increased attention as a viable strategy for making meaningful contributions to increases children's school-based PA participation (Norris et al., 2015; Martin & Murtagh, 2015), as well as to improvements children's classroom behavior (Goh et al., 2016; Grieco et al., 2016; Mahar et al., 2006) and academic performance (Donnelly & Lambourne, 2011; Kibbe et al.,2011; Mullender-Wijnsma et al., 2016; Resaland et al., 2018; Watson et al., 2017). However, little is known about the effect of

MI on children's MC, or whether classroom-based PA experiences support students' SEL. There is a dearth of programs developed specifically for the elementary classroom setting that specifically address children's MC and, therefore, limited information is available with respect to the feasibility of classroom-based initiatives to develop MC. Currently, although the CDC (2015) recommends MI as a key strategy to increase children's PA during school, few U.S. school districts (10.7%) require schools to provide MI (National Physical Activity Plan Alliance, 2018). Additionally, teachers and administrators perceive numerous barriers to using MI, such as lack of time, resources and training (Carlson et al., 2017; Dinkel et al., 2017; Dyrstad et al., 2018; Michael et al. 2019; Quarmby, Daly-Smith, & Kime, 2019; Webster et al., 2017).

Overall Purpose of Dissertation and Proposed Studies

In line with the focus of this introductory chapter and the aforementioned gaps in the related literature, the original purpose of this dissertation was to examine the potential of opportunities within the CSPAP framework, particularly beyond traditional PE classes, to bolster students' PE learning. However, given the onset of the COVID-19 pandemic during the development of this dissertation, parts of the originally proposed research needed to change accordingly. Three studies were proposed, including two systematic reviews and meta-analyses (Studies 1 and 2, respectively) and one school-based pilot intervention designed to increase elementary children's MC and SEL via MI (Study 3). COVID-19 prevented Study 3 from proceeding due to school closures. Therefore, instead of conducting an intervention, the author of this dissertation and his committee agreed to develop a research measure for assessing online PE instruction, based on the pronounced need for such a tool in the context of the pandemic. With this change, the overall purpose

of this dissertation was revised. Specifically, the updated purpose of this dissertation was to examine CSPAP-aligned strategies, particularly beyond traditional PE, for supporting students' PE learning. Using the CSPAP framework to categorize different approaches to PA interventions with elementary children, the purposes of Studies 1 and 2 were to conduct systematic reviews and meta-analysis of the effectiveness of such interventions in increasing children's MC and SEL, respectively. The purpose of Study 3 was to develop a systematic observation instrument for assessing synchronous online PE lessons.

CHAPTER 2

STUDY 1: SYSTEMATIC REVIEW AND META-ANALYSIS OF PHYSICAL ACTIVITY INTERVENTIONS TO INCREASE ELEMENTARY CHILDREN'S MOTOR COMPETENCE

INTRODUCTION

It is well established that physical activity (PA) is vitally important for the healthy growth and development of children (Centers for Disease Control and Prevention [CDC], 2018; World Health Organization [WHO], 2019); however, many children are not adequately active. Only 24% of youths from 5-17 years old in the United States are meeting the nationally recommended minimum of 60 minutes of PA per day (National Physical Activity Plan Alliance, 2018). Motor competence (MC) plays a major role in children's PA participation (Barnett et al., 2016; Fisher et al., 2005; Logan et al., 2015; Stodden et al., 2008; Robinson et al., 2015). MC can be defined as the capability to perform a wide range of motor acts or skills and involves both locomotor (e.g., running, jumping, and skipping) and object projection (e.g., throwing, catching, and kicking) skills (Haywood & Getchell, 2019). The development of MC during childhood is crucial for a healthy life since it will allow individuals to successfully participate in lifetime physical activities (Barnett et al., 2016; Robinson et al., 2015; Stodden et al., 2008). Children with greater MC were observed to spend more time in moderate-to-vigorous PA (Williams et al., 2008), whereas those with less developed MC appeared less physically active

(Barnett et al., 2016; Logan et al., 2015). Further, longitudinal evidence suggests that having higher levels of MC as a child is associated with being more physically active later in life (Barnett et al., 2009; Hulteen et al., 2018; Lloyd et al., 2014; Lopes et al., 2011). Conversely, low MC is hypothesized to result in decreased participation in PA in middle to late childhood, thus leading to a negative spiral of disengagement from an active lifestyle (Robinson et al., 2015; Stodden et al., 2008).

Developing children's and adolescents' MC is a primary goal of PE and is considered foundational to promoting lifetime participation in PA (Colvin, Markos, & Walker, 2016; Society of Health and Physical Educators [SHAPE] America, 2014). Developing a strong foundation in MC during early childhood can transition/apply to more specialized movement forms in organized games and sports (Gallahue et al., 2012; Utesch et al., 2019). For this to happen, regular involvement in context-specific and developmentally appropriate PA experiences is critical (Brian et al., 2020; Logan et al., 2012; Newell, 2020). However, focused programming to support children's MC development is decreasing for school-aged children, in tandem with a downward trend in the prevalence of PE (Trost et al., 2002; Tester, Ackland, & Houghton 2014).

A possible mechanism for extending students' motor skill learning beyond PE is a comprehensive school physical activity program (CSPAP; National Association for Sport and Physical Education [NASPE], 2008; CDC, 2019). The CSPAP model includes five components: (a) quality of PE, (b) PA during school, (c) PA before and after school, (d) staff involvement, and (e) family and community engagement (CDC, 2019). The CSPAP model has become a widely disseminated paradigm in literature concerning whole-of-school approaches to youth PA promotion (Webster et al., 2020). Through such an

approach, additional PA opportunities before, during, and/or after school are provided to (a) help students meet the nationally recommended 60 (or more) minutes of daily PA and (b) support students' learning of the knowledge and skills needed to participate in a lifetime of PA (NASPE, 2008). While researchers have established that school-related PA opportunities beyond PE can increase students' participation in PA (Pulling Kuhn et al., 2021; Erwin et al., 2013), less is known about the extent to which such opportunities may directly support children's development of MC. A clear understanding of the potential of different CSPAP components to enhance children's MC is lacking. The purpose of this study will be to conduct a systematic review and meta-analysis of the effectiveness of PA interventions in increasing the MC of elementary school children (5-12 years). The CSPAP framework will be used to consider the effectiveness of different intervention approaches in increasing children's MC.

METHOD

Registration and protocol

This study followed the updated Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Appendix A; Page et al., 2021) with additional recommendations for systematic meta-reviews (Hennessy et al., 2019) and was registered with the International Prospective Register of Systematic Reviews at https://www.crd.york.ac.uk/prospero/ (registration number CRD42020179866).

Inclusion/eligibility criteria

Studies with the following characteristics were included in our review:

1. Participants were aged 5-12 years (primary/elementary school);

2. PA interventions primarily focused on improving and assessing MC/FMS components;

3. Type of interventions: Any school-, home-, or community-based interventions for children with clear intent to improve MC/FMS proficiency;

4. Type of studies: Employed a Cluster-Randomized Controlled Trials (C-RCTs) design, RCTs, or rigorous (matched or statistically controlled) quasi-experimental design.

Exclusion criteria

Studies with the following characteristics were excluded from the review:

1. Studies that reported on population of focus outside of the age range defined above; or participants were not 'typically developing' (i.e. had a clinically diagnosed physical or intellectual disability or condition affecting movement, e.g. autism, visual impairment, cerebral palsy, traumatic brain injury/concussion);

2. Studies that did not aim to improve and assess at least one of MC/FMS components outcomes were excluded;

3. Studies reported as abstracts, theses/dissertations and unpublished literature were excluded.

Search strategy and terms

The studies were obtained on February 28, 2020 using seven electronic databases: PubMed/Medline, Embase, ERIC, SPORTDiscus, CINAHL, Web of Science, and PsycINFO. The search strategy consisted of four elements: study population (e.g., elementary school student), study design, intervention (e.g., PA and exercise), and

outcome measures (e.g., FMS and MC; Appendix B). The search was limited to peer-reviewed academic journal articles published in English in all available years.

Data extraction/collection process

Data were imported into Endnote X9.3 and duplicates were removed. The selected references were imported to a web□based software platform that streamlines the production of systematic reviews (Covidence systematic review software, Veritas Health Innovation, Melbourne, Australia, available at www.covidence.org). Two independent reviewers screened the titles and abstracts of retrieved records for possible inclusion. Of the records identified as possibly eligible, the full texts were obtained, and two independent reviewers assessed the records for inclusion. For each included study, two reviewers extracted data into a pre-defined Microsoft Excel (Microsoft Corporation, Redmond, WA, USA) data collection form. Data were extracted on the following: type of study design; the intervention approach, based on the CSPAP framework; the sample size; the intervention characteristics including session duration, frequency, length, delivery, and the name of programs; the types and methods of measured outcomes including the specific instrument; the fidelity of implementation measure; and the main results (Appendix C). For all steps in the screening process and data extraction, a third reviewer checked the data for errors, and discrepancies were resolved through discussion and consensus of judgement. If data were missing, authors were contacted.

Qualitative data synthesis

Extracted results showed information including the article reference, study design, intervention approach (i.e., CSPAP components used), study characteristics (country, school setting, school level, name of the intervention program, participants, intervention

deliverer, and MC outcomes, etc.), dose, main results, and implementation fidelity reporting. Results were organized into three sections by: (1) study design (i.e., C-RCT/RCT and N-RCT), (2) interventions addressing a single CSPAP component (i.e., PE, PADS, PABAS, and FCE) and (3) interventions addressing multiple CSPAP components (e.g., PE+1 additional component, PBAS+FCE).

Quantitative data synthesis

Effect sizes were calculated for the intervention group relative to the comparison group for each study. When the necessary data were not available in the original article, we requested it from the authors. If data could still not be obtained, we extracted the data from the graphs when available. If that was not possible, we excluded the study from the quantitative analysis. A meta-analysis for a given MC outcome was conducted if at least three studies reported interventions addressing the same CSPAP components and provided sufficient data for the calculation of effect size.

Pre- and post-intervention mean ± standard deviation (SD) for a given MC outcome, and sample size from each study were converted to Hedges' g effect size (Hedges, 1981). Specifically, we calculated standardized mean differences both for outcome scores at the end of the intervention period (i.e., post-intervention) and change-from-baseline (pre-intervention) outcomes. Scores post-intervention effect sizes refer to intervention group results compared with comparison or control group results after interventions. We did not include follow-up assessment data. In all analyses, we used the random-effects model to account for differences between studies that might impact the treatment effect (Deeks et al., 2008; Kontopantelis et al., 2013). The effect size values are presented alongside their respective 95% Confidence Intervals (CIs). Calculated effect

sizes were interpreted using the following scale: small (g<.40), moderate (g=.40–.07), and large (g>.07), according to the Cochrane Handbook (Higgins et al., 2019). Heterogeneity (i.e., between studies variability) was evaluated using the I-squared (I^2) statistic. I^2 values of <25%, 25-75%, and > 75% were considered to represent low, moderate, and high levels of heterogeneity, respectively (Huedo-Medina, Sanchez-Meca, Marin-Martinez, & Botella, 2006). The risk of bias was explored using the visual inspection of funnel plots and Egger's regression test (Egger, Smith, Schneider, & Minder, 1997). Publication bias was not produced as the meta-analyses included <10 studies/interventions (Higgins & Green, 2017).

A series of models were analyzed to address the following: (1) the pooled effect of PA interventions across all studies on elementary school age-children's MC (overall and by measurement), (2) the pooled effect of interventions using only PE compared to the pooled effects of other single-component interventions that did not use PE (PADS only, PABAS only, and PADS + PABAS + FCE) on children's MC, and (3) the pooled effects of interventions using PE plus additional CSPAP components (PE+1 and PE+2) on children's MC. In addition, moderation analyses were performed to explore the impact of potential explanatory variables and moderators (intervention duration [<6 months vs. ≥ 6 months], delivery agent [research team vs. school-based team vs. combined], and study design [C-RCT/RCT vs. N-RCT]) on the effect sizes with meta-regressions when sufficient data were available (i.e., at least ten studies for each explanatory variable; Higgins et al., 2019). The results were expressed as regression coefficients estimates, 95% CIs and the p value. All analyses were carried out using the Comprehensive Meta-

Analysis program (version 3.3.070; Biostat, Englewood, NJ, USA). The statistical significance threshold was set at $p<.05$.

Risk of bias assessment

Risk of bias in the included studies was assessed by two reviewers independently through discussion using the Cochrane Risk of Bias Tool (RoB 2.0) with additional considerations for C-RCTs and RCTs (Sterne et al., 2019), which consists of five domains and an overall judgment (Hennessy et al., 2019). The five domains are: 1) bias arising from the randomization process; 2) bias due to deviations from the intended interventions; 3) bias due to missing outcome data; 4) bias in measurement of the outcome; and 5) bias in selection of the reported result. Based on the answers (yes, probably yes, probably no, no, not applicable, no information) to a series of signaling questions in the guidance document, the judgment options within each domain consist of "low risk of bias," "some concerns", or "high risk of bias" (Sterne et al., 2019).

The N-RCT (i.e., quasi-experimental) studies were assessed with the Risk of Bias in Non-randomized Studies of Interventions tool (ROBINS-□; Sterne et al., 2016), which consists of seven domains and an overall judgement. The seven domains are: 1) bias due to confounding; 2) bias in selection of participants into the study; 3) bias in classification of interventions; 4) bias due to deviations from intended intervention; 5) bias due to missing data; 6) bias in measurement of outcomes and 7) bias in selection of the reported result (Sterne et al., 2016). Domain-specific risk of bias assessment was used to judge the overall risk of bias for each study. Disagreements between reviewers were resolved through discussion and consensus by a third evaluator. Before correcting for observed differences, the agreement between reviewers was assessed using a Kappa correlation for

risk of bias (κ > 0.8). A risk of bias graph was made via the *robvis* R package (McGuinness & Higgins, 2021).

RESULTS

A total of 6,064 search records were initially identified. The authors screened 3,804 records after removing duplicate records. This first level of screening, separated by title and abstract, identified 439 full-text articles to be reviewed for eligibility. Ultimately, of the remaining 286 articles, 27 studies were included in the qualitative synthesis and 26 studies were included in quantitative synthesis. The process of literature identification and selection is outlined in the PRISMA flowchart (Figure 2.1). The quality assessment for C-RCTs or RCTs revealed 5 studies as low risk in quality, 4 studies as having some concerns in quality, and 1 study as high risk (Figure 2.2). For N-RCTs, 9 studies were evaluated as low risk of bias, 6 studies as moderate quality, and 2 studies as serious risk of bias (Figure 2.3). Generally, the studies included a lack of clear description of randomization procedures and lack of clarity regarding drop-out rates. There are some studies that did not assess the fidelity of the interventions to determine if they were implemented as intended. Additionally, most of the studies had some concerns due to deviations from the intended interventions.

Characteristics of the included studies

Across the 27 studies, there were a total of 13,281 participants (49% female, 51% male) from 306 classes and 191 schools. The sample size ranged from 13 (Cliff et al., 2007) to 4,234 participants (Chen et al., 2016) with the age of intervention children ranging from 5 to 12 years. 10 studies were conducted in North America (i.e., Canada and United States), 7 in Europe (i.e., Finland, Germany, Ireland, Netherlands, Poland,

and United Kingdom), 8 in Australia, 1 in Asia (i.e., China), and 1 in South America (i.e., Brazil). Additionally, 4 studies were conducted in urban settings, 1 in rural and urban settings, 1 in a rural setting, and 2 in a suburban setting. The setting was not specified in 19 of the studies.

Descriptions of CSPAP components

Study design. 8 C-RCTs (30%), 2 RCTs (7%), and 17 N-RCT studies (63%) were included in this review.

C-RCT/RCT. 10 studies were C-RCTs or RCTs (Chan et al., 2019; Cohen et al., 2015; Laukkanen et al., 2015; Laurent et al., 2018; Lammle et al., 2016; Maskell et al., 2004; McKenzie et al., 1998; McWhannell et al., 2018; Miller et al., 2015, 2016), with the average number of schools and classes across all studies being 14 (range 1–91) and 33 (range 2–157), respectively (Lammle et al., 2016; Maskell et al., 2004). The average sample size was 369 students (range 28–1,736; Laurent et al., 2018; Lammle et al., 2016), and the total number of students was 3,054 (48% Female, 52% Male). The average intervention duration, frequency, and length was 30 weeks (range 5–96; Maskell et al., 2004; McKenzie et al., 1998), 2 times per week, and 55 minutes per session (range 15–120; Cohen et al., 2015; Laurent et al., 2018), respectively. For measurement of children's MC, 5 studies of 10 (50%) used Test for Gross Motor Development (TGMD-2nd Edition or -3rd Edition; Ulrich, 2000, 2019), 1 study (10%) used Körper-koordinationtest Für Kinder (KTK; Kiphard & Schilling, 2007), 2 studies used other measurements (e.g., Dordel-Koch-Test [DKT]; 20%), and 2 studies (11%) did not specify a measurement tool.

N-RCT. 17 studies were N-RCTs (i.e., quasi-experimental), with the average number of schools and classes across all studies being 4 (range 1–9; Chen et al., 2016; Lee et al., 2020a; Rudd et al., 2017; Silveria et al., 2018) and 6 (range 2–20; Rudd et al., 2016, 2017; Skowroński et al., 2019), respectively. The average sample size was 564 students (range 13–1,460; Burn et al., 2017a; Cliff et al., 2007), and the total number of students was 3,686 (49% Female, 51% Male). The average intervention duration, frequency, and length was 30 weeks (range 4–176; Okely et al., 2017; Platvoet et al., 2016), 2 times per week (range 1–5; Burrows et al., 2014; Cliff et al., 2007; Chen et al., 2016; Rudd et al., 2016), and 59 minutes per session (range 25–120; Bolger et al., 2019; Cliff et al., 2007; Okely et al., 2017; Rudd et al., 2016), respectively. For measurement of children's MC, 11 studies (64%) used TGMD-2 or -3, 4 studies (24%) used KTK, 2 studies used other measurements (e.g., PE Metrics; 11%), and 1 study (5%) did not specify a measurement tool. Two studies (Rudd et al., 2016, 2017) used TGMD-2 and KTK.

Single CSPAP component interventions.

Considering the CSPAP framework, 9 studies out of 18 (50%) used only PE for the intervention, 3 studies (17%) used only PADS, 4 studies (22%) used only PABAS, and 2 studies (11%) used only FCE. No study used only SI for the intervention.

PE. For the studies that used only PE as the intervention approach, the average number of schools and classes across all studies were 4 (range 1–9; Chen et al., 2016; Maskell et al., 2004; Rudd et al., 2017) and 7 (range 2–20; Maskell et al., 2004; Rudd et al., 2016, 2017), respectively. The average sample size was 691 students (range 42–4,234 (Chen et al., 2016; Maskell et al., 2004), and the total number of students was 3,054 (50%

Female, 50% Male). The average intervention duration, frequency, and length was 20 weeks (range 4–96; Chen et al., 2016; Platvoet et al., 2016), 2 times per week, and 57 minutes per session (range 25–120; Bolger et al., 2019; Rudd et al., 2016), respectively.

PE interventions involved PE lessons that incorporated revised fundamental motor skill (FMS) activities (Bolger et al., 2019); movement activities related to specific motor skills (Chan et al., 2019); a PE curriculum that included motor skill themes and physical fitness activities (Chen et al., 2016); goal-directed learning (Gu et al., 2017; Platvoet et al., 2016); a movement program (Brain Gym) involving a series of simple-to-challenging FMS intended to enhance cognitive processing, psychomotor and whole-brain learning (Maskell et al., 2004); the Professional Learning for Understanding Games Education (PLUNGE) program, which aimed to increase the complexity of challenges experienced through gameplay-situated learning for the improvement of FMS (Miller et al., 2015); and a gymnastics curriculum developed by Gymnastics Australia, which aimed to develop stability, locomotor and object control skills, and general body coordination (Rudd et al., 2016, 2017). The intervention deliverer varied across interventions. 1 study (11%) was delivered by a research team, 7 studies (78%) by a school-based team (i.e., Trained PE teachers, classroom teachers and students) and 1 study (11%) by a combined team (e.g., research team, school-based team, and parents). Additionally, 4 studies (44%) reported fidelity of intervention using observation and/or checklists.

PADS. For the studies that used PADS as the single intervention component (Johnstone et al., 2017; Lammle et al., 2016; Okely et al., 2017), the average number of schools and classes across all studies being 35 (range 7–91) and 84 (range 11–157, respectively (Johnstone et al., 2017; Lammle et al., 2016). The average sample size was

826 students (range 336–1,736; Johnstone et al., 2017; Lammle et al., 2016), and the total number of students was 2,479 (55% Female, 45% Male). The average intervention duration, frequency, and length was 81 weeks (range 20–176 (Johnstone et al., 2017; Okely et al., 2017), 2 times per week, and 33 minutes per session (range 15–60 (Johnstone et al., 2017; Lammle et al., 2016), respectively. PADS interventions involved structured games to increase children's FMS (Johnstone et al., 2017); short daily classroom exercises (Lammle et al., 2016); and a whole-of-school health promotion approach aimed to develop children’s FMS by modifying the physical and social environment (Okely et al., 2017). 2 studies (75%) were delivered by a school-based team and 1 study (25%) by a combined team. None of the studies reported fidelity of intervention.

PABAS. For studies that used only PABAS as an intervention approach, the average number of schools and classes across all studies was 7 (range 1–16 ;Lee et al., 2020a; McWhannell et al., 2018) and 3 (Lee et al., 2020b), respectively. The average sample size was 63 students (range 31–146; Lee et al., 2020a; McWhannell et al., 2018), and the total number of students was 252 (60% Female, 40% Male). The average intervention duration, frequency, and length was 13 weeks (range 8–26; Lee et al., 2020a; McWhannell et al., 2018), 3 times per week (range 2–5; Burrows et al., 2014; McWhannell et al., 2018), and 60 minutes per session, respectively.

The PABAS interventions involved outdoor low-organized games and indoor sports-based activities including swimming, floor hockey, and soccer as after school activities (Burrows et al., 2014); an after school program aiming to teach children the 12 basic motor skills from the TGMD-2 criteria (Lee et al., 2020a, 2020b); and an after

school club program that included multi-games activities, which focused on FMS development by using offering numerous opportunities for practice with learning cues (McWhannell et al., 2018). 3 studies (75%) were delivered by a research team and 1 study (25%) by an after school-based team (i.e., after school program leaders). Only 1 study (25%) reported fidelity of intervention (using field observations).

FCE. For studies that used FCE as the single intervention component, the average number of schools and classes across studies was 1 and 2, respectively. The average sample size was 193 students, and the total number of students was 385 (49% Female, 51% Male). The average intervention duration, frequency, and length was 27 weeks (range 6–48), 2 times per week, and 60 minutes per session, respectively (Laukkanen et al., 2015; Laurent et al., 2018).

The FCE interventions involved family involvement by providing tailored counseling (Laukkanen et al., 2015), structured PA homework/materials, educating parents to an increase children's MC, and goal-setting (Laukkanen et al., 2015; Laurent et al., 2018). 1 study (50%) was delivered by a research team (i.e., coaches and research assistants) and 1 study (50%) by a combined team (i.e., research team and parents). Both studies (100%) reported fidelity of intervention using observation and checklists.

Multiple CSPAP components interventions.

A total of 9 studies (33%) used intervention approaches that could be mapped onto multiple components within the CSPAP framework. The most commonly used components in multicomponent approaches were PE (n=8) followed by SI (n=7) and PADS (n=5). PABAS (n=3) and FCE (n=3) were included in less than half of the

multicomponent studies. One multicomponent intervention did not include a PE component (Cliff et al., 2007). No study included all five CSPAP components.

PE + 1 additional CSPAP component. 3 of the studies (33%) reported an intervention that included PE and 1 additional CSPAP component. 2 studies included SI and 1 study included PABAS. The average number of schools and classes across all studies were 3 (range 1–7; McKenzie et al., 1998; Miller et al., 2016) and 21 (range 2–56; McKenzie et al., 1998; Skowroński et al., 2019), respectively. The average sample size was 202 students (range 31–467; McKenzie et al., 1998; Skowroński et al., 2019), and the total number of students was 605 (29% Female, 71% Male). The average intervention duration, frequency, and length was 50 weeks (range 6–96; McKenzie et al., 1998; Miller et al., 2016), 3 times per week (range 1–4; Miller et al., 2016; Skowroński et al., 2019), and 45 minutes per session (range 30–60; McKenzie et al., 1998; Miller et al., 2016), respectively.

PE+1 interventions involved the Professional Learning for Understanding Games Education (PLUNGE) program that aimed to improve children's FMS in PE lessons through a professional learning process involving classroom teacher education and mentoring (Miller et al., 2016); a Sports, Play, and Active Recreation for Kids (SPARK) based PE program designed to enhance children's motor skills through a classroom teacher professional development program (McKenzie et al., 1998); and PE lessons combined with an extracurricular after school program (From Fun To Sport) with an emphasis on the development of children's FMS (Skowroński et al., 2019). All three interventions were delivered by a school-based team (i.e., PE teachers and trained

classroom teachers). Only 1 study (33%) reported fidelity of intervention using lesson observations (Miller et al., 2016).

PE + 2 additional CSPAP components. 3 studies (33%) reported interventions that included PE and 2 additional CSPAP components. 2 studies included the combination of SI and PADS with PE (Burn et al., 2017a; Nathan et al., 2017), and 1 study included SI and FCE with PE (Silveria et al., 2018).

The average number of schools and classes across all studies was 2 and 3, respectively. The average sample size was 664 students (range 174–1460; Burn et al., 2017a; Nathan et al., 2017), and the total number of students was 1991 (50% Female, 50% Male). The average intervention duration, frequency, and length was 11 weeks (range 10–12), 2 times per week, and 50 minutes per session (range 30–60), respectively Burn et al., 2017a; Nathan et al., 2017; Silveria et al., 2018).

The PE+2 interventions involved a CSPAP-based gross motor skill development program including the Dynamic PE for Elementary School Children curriculum during PE lessons, PA engagement opportunities throughout the school day during recess and regular classroom time (during which teachers integrated PA into academic lessons and classroom activity breaks via stretching, walking, jumping, or relaxation activities), and SI that provided teacher professional training to increase the quality of PE (Burn et al., 2017a); the Great Leaders Active StudentS (GLASS) program that included trained students who instructed their peers to improve FMS during PE lessons and classroom settings, and trained teachers supporting their peers' instruction, which contributed an SI component to the program (Nathan et al., 2017); and physical exercise sessions during PE lessons (e.g., circuit training, aerobic/sports activities, and recreational games, etc.),

parent support to promote PA during after school classes, and nutritional education sessions (e.g., goal setting and dietary counselling with parents; Silveria et al., 2018). 1 study (25%) was delivered by a school-based team (i.e., PE teachers, classroom teachers, and students; Nathan et al., 2017) and 2 studies (75%) by a combined team (i.e., PE teachers, classroom teachers, medical or healthcare staff, parents, PA leaders, and the research team; Burn et al., 2017a; Silveria et al., 2018). Only 1 study (33%) reported fidelity of intervention using observations and a checklist (Nathan et al., 2017).

PE + 3 additional CSPAP components. 2 studies (22%) reported interventions that included PE and 3 additional CSPAP components. 1 study included the combination of SI, PADS, and FCE with PE and 1 study included SI, PADS, and PABAS with PE. The average number of schools and classes across all studies was 7 and 25, respectively. The average sample size was 667 students (range 357–976; Burn et al., 2017b; Cohen et al., 2015), and the total number of students was 1333 (44% Female, 56% Male). The average intervention duration, frequency, and length was 42 weeks (range 36–48), 3 times per week (range 1–5), and 98 minutes per session (range 75–120), respectively (Burn et al., 2017b; Cohen et al., 2015).

One PE+3 intervention involved a CSPAP-based program that aimed to optimize the quality of PE. The intervention provided PA opportunities before and after school as well as during recess/lunch time, which created a number of opportunities for children to engage in free play or semi-structured PA by applying skills learned during PE lessons. Additionally, PA was integrated into academic lessons and classroom activities, and SI was addressed with continuous teacher training and assistance throughout the intervention (Burn et al., 2017b). The other PE+3 intervention involved the

implementation of six PA policies to support the promotion of PA and FMS competency within the PE lessons in combination with SI and PADS through teacher professional learning, student leadership workshops, and PA promotion tasks to achieve awards during recess and lunch. In addition, the intervention incorporated FCE via school–community connections (e.g., inviting local sporting organizations to assist with school sport programs) as well as a range of approaches targeting the home environment (e.g., newsletters, parent evening, and FMS homework; Cohen et al., 2015). Both studies were delivered by a combined team (i.e., PE teachers, classroom teachers, principals, parents, PA leaders, and community leaders; Burn et al., 2017b; Cohen et al., 2018) and both reported the fidelity of the intervention using observations and a checklist (Burn et al., 2017b; Cohen et al., 2015).

Multicomponent interventions without PE. 1 study (11%) included the combination of PABAS and FCE with no PE component (Cliff et al., 2007). The total number of students was 13 (62% Female, 38% Male). The intervention duration, frequency, and length were 10 weeks, 1 time per week, and 120 minutes per session, respectively. The intervention program involved a community-based program with an additional home-based PA motor development program using goal-setting and parental motivation strategies (Cliff et al., 2007). The intervention was delivered by a combined team (i.e., researchers and parents). The study did not report fidelity of intervention.

Meta-Analysis

Effectiveness across all interventions. The meta-analysis for total 26 studies indicated a statistically significant and large pooled intervention effect on children's total MC (Hedge's g = 0.71, 95% CI=0.60–0.81; p=.000; I^2 = 78.4%; Figure 2.4). The relative

weight of each study in the analysis ranged from 1.40 to 6.22%. For all included studies, Egger's regression test for asymmetry of the funnel plot was not significant (β = 0.32, p=0.17), indicating no evidence of publication bias (Figure 2.5). Results from the meta-regression found that intervention duration (β = -0.04; 95% CI=-0.29–0.19; p=0.69), delivery agent (β = -0.22; 95% CI=-0.65–0.21; p=0.31), and study design (β = 0.13; 95% CI=-0.10–0.36; p=0.28) were not found to be a statistically significant moderator variables/factors affecting overall study effect sizes (i.e., children's total MC).

In a subsequent analysis, the studies adopting the TGMD -2 or -3 tests were compared to studies including other types of assessments. The latter analysis was conducted as a proxy for effects of types of MC measurements. 17 studies measured children's MC using the TGMD-2 or -3 tool. The meta-analysis for studies indicated a statistically significant and large pooled intervention effect on children's total MC (Hedge's g = 0.79, 95% CI=0.63–0.95; p=.000; I^2 = 38.2%; Figure 2.6). The relative weight of each study in the analysis ranged from 3.88 to 9.30%. Additionally, 11 studies measured children's MC using other measurement tools (e.g., KTK, PE Metrics, and DKT). The meta-analysis for these studies indicated a statistically significant and moderate pooled intervention effect on children's total MC (Hedge's g = 0.57, 95% CI=0.42–0.72; p=.000; I^2 = 64.1%; Figure 2.7). The relative weight of each study in the analysis ranged from 3.77 to 12.62%. Specifically, 5 studies measured children's MC using KTK tool. The meta-analysis for studies indicated a statistically significant and moderate pooled intervention effect on children's total MC (Hedge's g = 0.41, 95% CI=0.28–0.57; p=.000; I^2 = 86.7%; Figure 2.8). The relative weight of each study in the analysis ranged from 4.04 to 12.82%.

Effectiveness of PE only vs. other single component interventions. 9 studies used only PE as the intervention approach. The meta-analysis for these studies indicated a statistically significant and large pooled intervention effect on children's total MC (Hedge's g = 0.79, 95% CI=0.55–1.04; p=.000; I^2 = 66.7%; Figure 2.9). The relative weight of each study in the analysis ranged from 7.40 to 13.30%. 8 other studies used non-PE single component intervention approaches (i.e., PADS+PABAS+FCE). The meta-analysis for these studies indicated a statistically significant and moderate pooled intervention effect on children's total MC (Hedge's g = 0.48, 95% CI=0.29–0.68; p=.008; I^2 = 85.0%; Figure 2.10). The relative weight of each study in the analysis ranged from 10.81 to 24.62%.

In a subsequent analysis, the studies that used non-PE single component interventions were analyzed. Specifically, 3 studies used only PADS as the intervention approach. The meta-analysis for these studies indicated a statistically significant and moderate pooled intervention effect on children's total MC (Hedge's g = 0.48, 95% CI=0.27–0.69; p=.000; I^2 = 76.5%; Figure 2.11). The relative weight of each study in the analysis ranged from 14.20 to 45.38%. Additionally, 3 studies used only PABAS as the intervention approach. The meta-analysis for these studies indicated a statistically significant and moderate pooled intervention effect on children's total MC (Hedge's g =0.50, 95% CI=0.13–0.89; p=.008; I^2 = 0.0%; Figure 2.12). The relative weight of each study in the analysis ranged from 25.28 to 46.18%.

Effectiveness of interventions addressing multiple CSPAP components. 3 studies reported intervention approaches that used PE and one additional CSPAP component (PE +1) to increase children's MC. The meta-analysis for these studies

indicated a statistically significant and moderate pooled intervention effect on children's total MC (Hedge's g = 0.64, 95% CI=0.33–0.95; p=.000; I^2 = 48.2%; Figure 2.13). The relative weight of each study in the analysis ranged from 13.25 to 18.30%. Additionally, 3 studies reported intervention approaches that used PE and two additional CSPAP components (PE +2) to increase children's MC. The meta-analysis for these studies indicated a statistically significant and moderate pooled intervention effect on children's total MC (Hedge's g = 0.55, 95% CI=0.27–0.82; p=.000; I^2 = 52.9%; Figure 2.14). The relative weight of each study in the analysis ranged from 12.98 to 32.75%.

DISCUSSION

The purpose of this systematic review and meta-analysis was to synthesize the evidence of the effectiveness of PA interventions in increasing MC as a primary outcome of children aged 5-12. Twenty-seven studies met the inclusion criteria and were included in the qualitative and 26 studies were included in the quantitative analysis. The CSPAP framework was used to categorize the different intervention approaches. The results of the aggregate meta-analysis indicate that effect sizes for the development of MC from pre-post intervention ranged from small to large. In light of our results, a wide range of CSPAP-aligned PA intervention approaches appear to be promising avenues in enhancing children's MC. However, there is considerable variation in study design, sample size, delivery agent, and study context, and studies were implemented in over 11 countries across diverse settings. Additionally, the results cannot show clear evidence that increased PA duration or frequency (i.e., dose) has a detrimental effect on the development of children's MC, which aligns with McDonough, Liu & Gao (2020).

Results of this review indicated that the majority of studies included PE as a component of either a single (33%) or multicomponent (30%) approach and showed beneficial effects on the development of children's MC. Specifically, PE commonly included the integration of movement activities with cognitively challenging PA learning experiences related to FMS and implementation of an established curriculum (e.g., SPARK) with professional teacher training (Chen et al., 2016; McKenzie et al., 1998; Miller et al., 2016). Especially in PA interventions that employ complex, challenging learning tasks, how such activities are delivered and implemented may crucially affect learning outcomes (Pesce et al., 2019). Jiménez-Díaz et al. (2019), in a review of 36 articles, present that naturally occurring PE classes were less effective at increasing children's MC than a research specialist-led motor intervention, based on PE teachers' lack of expertise for designing and implementing developmentally appropriate movement activities. However, our results showed that school-based teams (i.e. PE teachers and classroom teachers) can play a crucial role in increasing children MC with professional training and structured curriculum. Likewise, ongoing teacher training and support appears to be a key element of effective PE curriculums and successful interventions by enhancing the unique features of qualitative enrichment (Armour et al., 2017; Hastie et al., 2015; Tannehill et al., 2021).

The PE-based programs most often evaluated outcomes related to PA, fitness, and body composition (Errisuriz et al., 2018). Conversely, most of the included studies focused on the development of MC beyond PA opportunities. These results are consistent with those of a previous systematic review that found that FMS-based intervention programs appeared to have larger effects than interventions focused strictly on increasing

PA (Collins et al., 2019). Further, when it comes to curriculum, research has demonstrated and noted the importance of structure when promoting children's motor skill development (Logan et al., 2012; Robinson & Goodway, 2009). In this review, a number of the PE components within CSPAP framework assessed an enhanced PE curriculum with a focus on optimal MC development as compared to traditional PE or free play (e.g., Chan et al., 2019; Miller et al., 2015), some simply tested the benefit of an additional time allotment of PE (Skowroński et al., 2019), and some compared both modified PE and time spent in PE lessons (Gu et al., 2017; Platvoet et al. (2016)). Overall, implementing a purposefully designed intervention approach with PE lessons had a positive effect on the development of MC. Further, of those studies that compared PE intervention programs versus typical PE (Okely et al., 2017), results support the importance of the quality of instructional approaches that enable students to have developmentally appropriate tasks/activities with learning cues, multiple opportunities for individual practices in a mastery climate, and individualized feedback (Nesbitt, Fisher, & Stodden, 2021; Palmer, Chinn, & Robinson, 2017; Rink, 2020; Ward, 2020). These results were quite similar to those demonstrated by Morgan et al. (2013), who highlighted the benefits of using a pedagogical approach to develop children's FMS in PE.

In this review, we considered dose (i.e., as the amount of time/duration devoted to motor skill instruction and practice; Robinson, Palmer, & Meehan, 2017), specifically <6 months vs. ≥6months, as a possible mediating factor in the effectiveness of PA interventions on MC development in children. Based on the results, however, the intervention dosage needed to obtain MC proficiency is unclear. For instance, some studies report significant improvements in children's MC after a 550-min dose over 13

weeks (Chan et al., 2019), 1,400 min dose over 8 weeks (Lee et al., 2020a), 1,400 min dose over 12 months (Silveira et al., 2018), and 2400 min dose over 20 weeks (Johnstone et al., 2017), whereas other studies fail to see significant effects after a 480 (Maskell et al., 2004) or 3600-min dose (Laukkanen et al., 2015) at 5 weeks and 12 months, respectively. Similarly, previous literature demonstrates inconsistencies regarding the amount of intervention needed to produce positive developmental changes in MC. Wick et al. (2017) found interventions conducted from 1 to 5-months had a larger effect on FMS than interventions lasting over 6 months. In addition, a recent meta-analysis study indicated that children aged 3-5 need to practice their FMS with a teacher-led intervention regularly (i.e., 3 times per week for ≤ 6 months) to achieve significant improvement in MC (Engel et al., 2018). Specifically, Van Capelle et al. (2017) suggested that interventions for increasing FMS must be implemented more than 3 times per week and that sessions should last longer than 30 minutes. However, the meta-analysis by Logan et al. (2012) reported a nonsignificant relationship between effect sizes of FMS improvements and intervention duration with a dosage between 500 and 1,400 min. A possible explanation for the results is that there was heterogeneity in study length across the included studies which ranged from 4 weeks to 192 weeks and was administered 1–5 times per week in a 15–120 minutes per session. Another possible explanation is to speculate the “ceiling effect” in which children might already achieve better performance in the early stages of the intervention. As a result, more time (quantitative aspect) may not necessarily translate to better performance (quality aspect). Robinson et al. (2017) presented that as little as 600 min of high-quality instruction during the intervention program can significantly improve children’s MC. Thus, future

research is warranted. It would be beneficial to examine the impact that different intervention dosages (e.g., duration and frequency) would have on children's MC development effectively under similar PA intervention conditions. Additionally, most studies did not report the dose received (i.e., on-task time in the tasks/activities), which is an important area for future research because motor skill development theory shows that one of the key factors is the number of correct practice trials a child completes (Silverman, 1991). Ultimately, understanding patterns of change resulting from different ranges of intervention dosages could illustrate how only minimal amounts of time could lead to positive developmental changes in MC and help establish recommendations and policies for practitioners implementing CSPAPs.

Interestingly, studies mainly reported multiple component interventions, specifically addressing the CSPAP components FCE and SI (Burn et al., 2017a, 2017b; Cliff et al., 2007; Cohen et al., 2015; McKenzie et al., 1998; Miller et al., 2016; Nathan et al., 2017; Silveria et al., 2018). The multicomponent interventions were collaboratively delivered by a variety of facilitators, such as PE teachers, classroom teachers, administrators, coaches, community leaders, parents, and medical or healthcare staff. The results of this review parallel previous interventions that involve parents as promoters for PA and MC in their own children (Dozier et al., 2020; Gustafson & Rhodes, 2006). Overall, it seems reasonable to assert that the FCE and SI components of the CSPAP framework function as important elements in the support system for PA program implementation in schools (Webster et al., 2020), and can help to enhance children's MC. However, there is insufficient evidence specific to each component (FCE or SI) to make conclusions about the its specific contribution to MC, and further research is needed to

determine which strategies are most effective for optimizing FCE and SI to support the development of MC in children.

Overall, there has been a lack of variety in theories used to guide intervention development. The studies in this review used the socio-ecological model (Cohen et al., 2015; Silveira et al., 2018), motivation theory (Cliff et al., 2007; Gu et al., 2017), and social cognitive theory (Laukkanen et al., 2015). However, other theoretical perspectives should be considered, as well. For example, interventions could incorporate strategies such as encouraging teachers to provide positive feedback and emphasizing mastery of skills rather than competition. These practical strategies reflect constructs related to motivation theories such as Self-Determination Theory (Deci & Ryan, 1985) and Achievement Goal Theory (Ames & Archer, 1988). Moreover, there is a lack of strong process evaluation across studies. While intervention programs demonstrated improvements in children's MC, multiple components were typically implemented simultaneously. Nearly half of the studies did not measure intervention implementation elements (e.g., fidelity and selection of participants). Additionally, some studies did not describe their instructional strategies in detail. As such, it is unclear how these variables affected study findings. It is important to show the study context and resources to improve the interpretation of research findings.

The meta-analysis results indicated that the pooled effect sizes of all CSPAP-aligned interventions to increase children's overall MC were statistically significant, with 11 studies (42%) reporting large effect sizes. However, there were a small number of heterogeneous studies included in the meta-analysis. Subsequently, a subgroup comparison between measurements was performed. Studies were separated by

measurement tool (i.e., TGMD-2 or-3 vs. other measurements); the studies that used TGMD-2 or -3 had a large effect on MC (e.g., Johnstone et al., 2017; Lee et al., 2020a, 2020b), whereas other assessments had a moderate effect on MC (Hedge's g = 0.79 vs. 0.57). Thus, not all CSPAP-aligned PA intervention programs have the same effect on the development of MC. For example, Rudd et al., (2016, 2017), which assessed changes in both TGMD-2 and KTK, reported different effect sizes (Hedge's g = 0.79 vs. 0.41) in the development of MC in their gymnastics intervention group. Since studies using TGMD-2 or -3 often show large effect sizes, they could thus be an effective way to assess the development of children's MC (as opposed to other measurements). Although these measurements are commonly used for assessing children's MC, the variety in scoring criteria protocol might provide different aspects of MC across the different movement dimensions evaluated (Logan et al., 2012; Rudd et al., 2016). Research has also shown that there is a low-to-moderate correlation between TGMD-2 and KTK in children (Ré et al., 2018). Therefore, further studies on the effect of CSPAP-aligned PA interventions on MC should carefully consider the types of assessments and their associations with intervention outcomes.

The current meta-analysis found that single PE component interventions had a larger effect on MC than other single-component approaches. A common misconception is that MC development is a naturally occurring phenomenon; however, the literature suggests that it must "be practiced, taught, and reinforced through developmentally appropriate movement programs" (Robinson & Goodway, 2009, p. 533). A large number of previous studies using the CSPAP framework focus mainly on the potential of PE to provide enough amounts of PA, that is, its contribution to the achievement of daily PA

recommendations (Webster et al., 2020). The main results of the meta-analysis showed that the PE component is foundational to learning and developing MC in children (Lopes, Stodden, & Rodrigues, 2017; Webster et al., 2020). Previous research found that PE contributed to improving elementary school students' manipulative skills (Chen et al., 2016) and motor skill competence (Lopes et al., 2017). In a meta-analysis study, Dudley et al. (2011) also presented that PE can be efficacious in improving MC in primary school children. Moreover, the quality of instruction and time spent in practice are of utmost importance in improving MC (Gallahue et al., 2012). However, limited research is devoted to studying the unique potential of PE within the CSPAP framework to develop MC in schoolchildren aged 5-12. Additionally, our results suggest that given the limited PE curriculum time in elementary schools, strategies to engage classroom teachers and/or parents in both school-based lessons and to support practice opportunities outside of PE class and school may be a worthwhile target for future interventions. That is, implementing a PA program using other single components in the CSPAP framework has the potential to support PE's goal of developing children's MC.

There was minimal effectiveness of adding other CSPAP components to PE for the development of children's MC. Many of the interventions included in the current review were multicomponent interventions. We expected that multiple components being used to increase MC may be a more effective approach than the single PE component approach. Unfortunately, the results suggest multicomponent interventions (adding other components to PE) have had minimal impact on the development of children's MC as indicated by the effect size values. A possible explanation for our results is that the quality of PA programs might be more important than the quantity of CSPAP

components used to increase children's MC. More research, therefore, is warranted to examine how the quality of PA experiences provided through each CSPAP component can impact the MC of school-age children.

In this review, moderation analyses were performed to explore the effectiveness of potential moderators (i.e., study length, delivery agent, and study design) on the effect size of MC with meta-regression (Hennessy et al., 2019). In general, our findings suggest that the effects of potential moderator variables had no significance on children's MC. In line with the present results, Lorås (2020) also showed that participants' ages, the total amount of time for intervention, and the type of MC measurement were not statistically significant moderators of effect size. However, a handful of studies had insignificant heterogeneity, indicating substantial differences in study contexts and characteristics. Additionally, there was limited information available for some moderator variable analyses. Thus, further studies should collect and report more complete data so that potential moderators can be examined to help us better understand the effects of potential moderator variables on the development of MC.

Strengths and limitations

To the best of our knowledge, this is the first review to examine CSPAP-aligned PA intervention effects on the MC of children aged 5-12 years. Although previous meta-analyses on this topic have been conducted, they were limited to FMS in the early childhood age band (Engel et al., 2018; Jones et al., 2020; Van Capelle et al., 2017; Wick et al., 2017). In this review, we included elementary school age and, not only FMS, but also motor coordination and motor proficiency, representing a broader range of MC outcomes. Additionally, our review provided insight into different intervention

approaches to change children's MC within the context of the CSPAP framework, highlighting the quality of PE as a particularly effective foundational component. Another strength of this review was that we used a sensitive search strategy to ensure relevant studies were not missed. In addition, a rigorous review methodology, including independent, duplicate reviews of selected studies, ensured most studies were captured (Hennessy et al., 2019).

This study also has several limitations. First, for practical reasons, we only included peer-reviewed studies published in English; non-English publications, and therefore further comparative evidence, may have been available on the topic. Second, there was noticeable heterogeneity of study approaches and assessment tools used to test children's MC across studies, which makes it difficult to compare findings across the studies. Nevertheless, validated testing instruments were utilized across included studies which minimized a major domain of bias and further strengthened the overall evidence of this review. Finally, the review included small sample sizes and some articles were feasibility studies or pilot trials.

CONCLUSION

The current review provides a unique contribution to the literature through its primary focus on the effectiveness of CSPAP-aligned PA interventions on children's MC. In light of our results, single and/or multi component intervention approaches, within the CSPAP framework, appear to be a promising avenue to promote MC in school-aged children (5-12 years old). This review also highlights that CSPAP-aligned PA programs should be tailored to the context within which they are delivered, most notably the PE component which can be best adapted to the context through professional teacher

training. Additionally, combining different PA intervention strategies (e.g., goal setting and reinforcement) with the SI/FCE components should be considered to improve MC through increased engagement and motivation.

Beyond these findings, this study identified avenues for future research. To increase intervention engagement and efficacy, future studies should examine the impact of a greater emphasis on children's MC. It is important to improve our understanding of which CSPAP-aligned PA intervention approaches are more effective than others by stratifying for the target groups, the setting, and the characteristics of the interventions. In other words, we need to identify and investigate well-designed interventions, including tailored types of PA programs. Additionally, future research should use stronger methodological approaches and consider expanding the theoretical ground for this research. Specifically, in order to make intervention studies more robust, research using high-standard randomization procedures to investigate the ability of CSPAPs to improve children's MC is needed. Ultimately, we should pursue how to effectively translate the evidence into practice to better conceptualize CSPAPs designed for children's MC development.

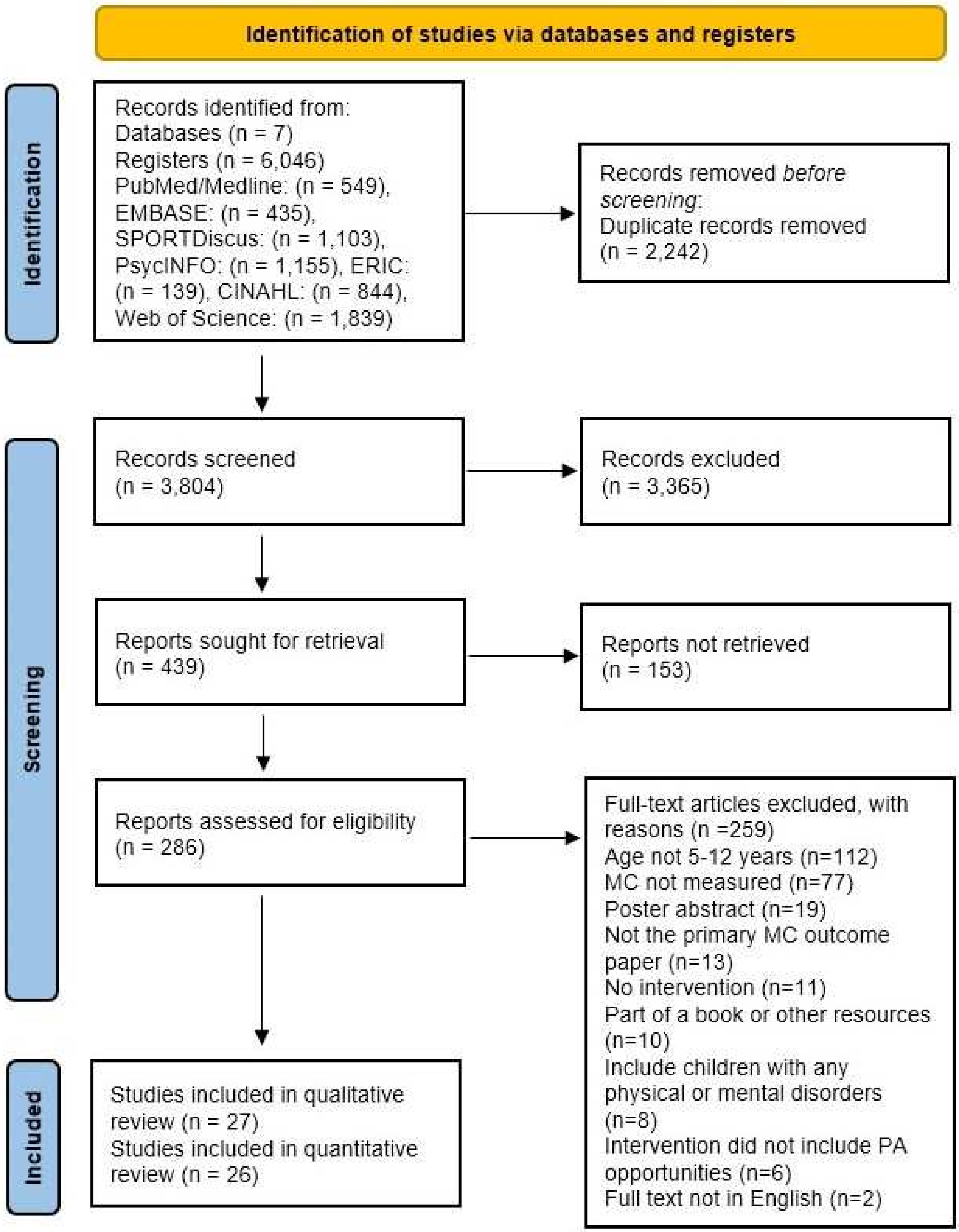

Figure 2.1. PRISMA flow chart of the search process of screened, included, and excluded articles.

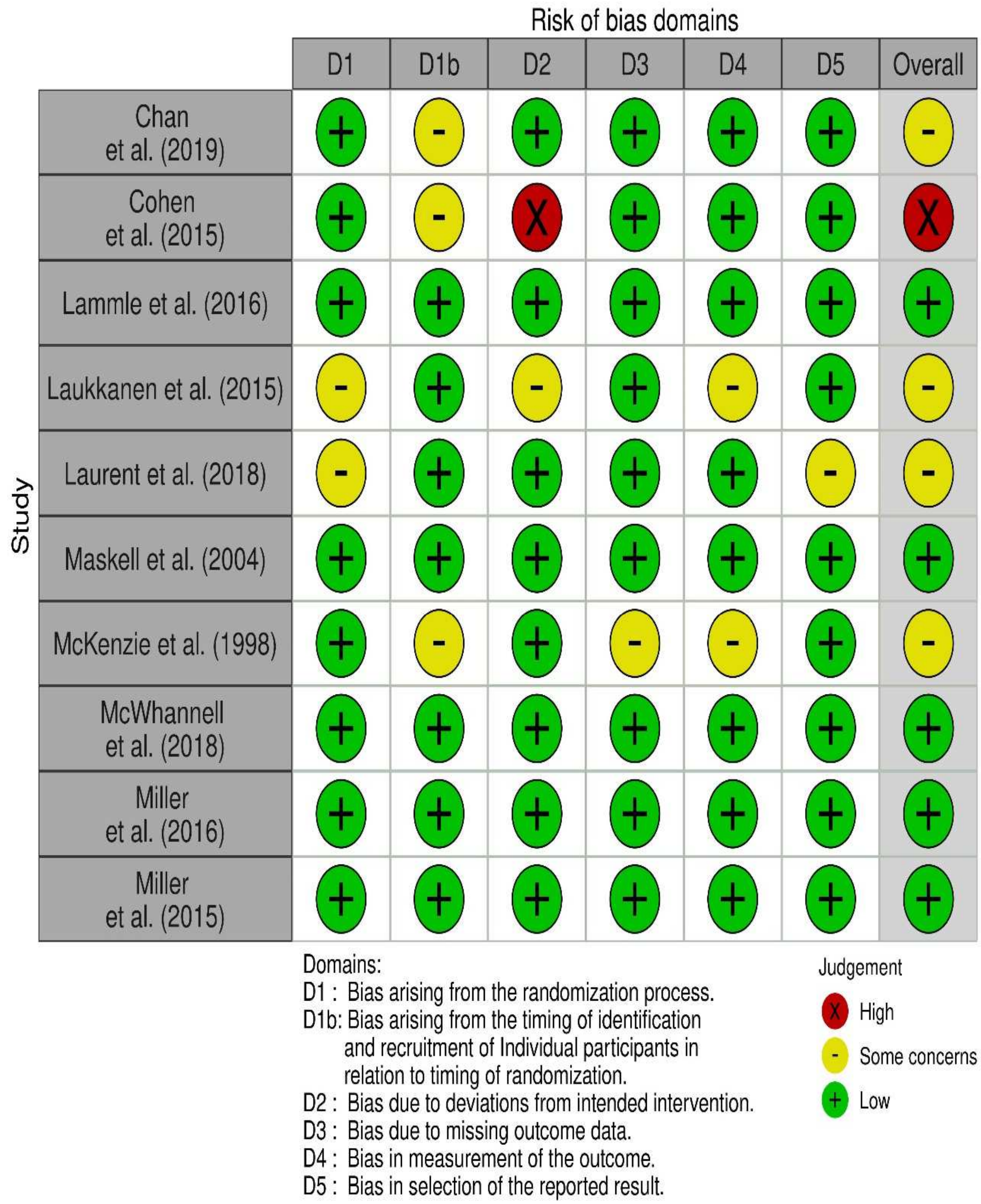

Figure 2.2. Quality assessment for C-RCT or RCT studies based on Revised Cochrane Risk of Bias tool for randomized trials (RoB 2.0).

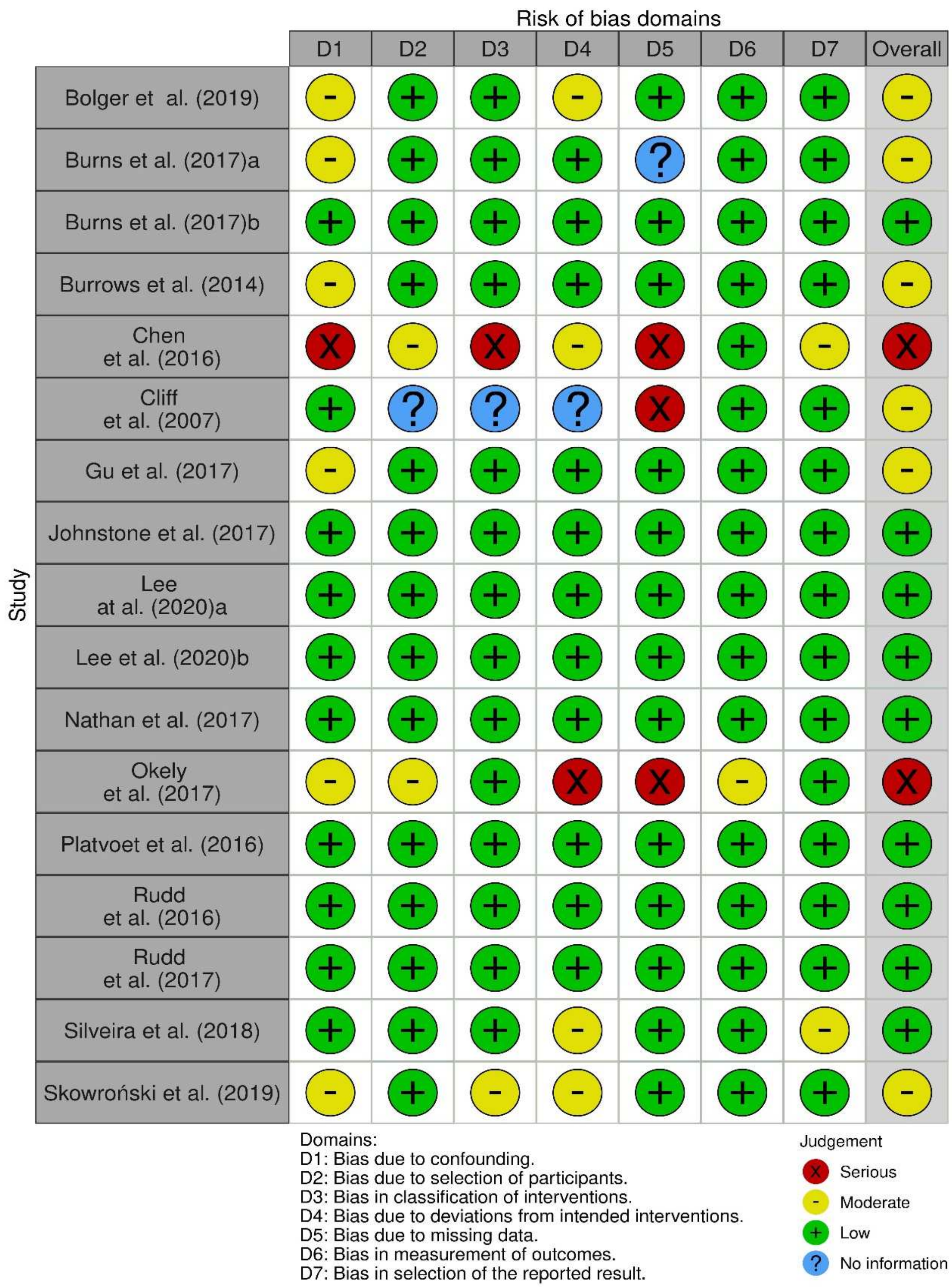

Figure 2.3. Quality assessment for N-RCT studies based on Risk of Bias in Non-randomized Studies of Interventions tool (ROBINS-I).

Study name	Hedges's g	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
Bolger et al. (2019)	1.929	0.177	0.031	1.583	2.276	10.923	0.000
Burn et al. (2017)a	0.421	0.027	0.001	0.367	0.474	15.419	0.000
Burn et al. (2017)b	0.595	0.035	0.001	0.527	0.664	17.005	0.000
Burrows et al. (2014)	0.349	0.358	0.128	-0.353	1.051	0.975	0.329
Chan et al. (2019)	0.685	0.124	0.015	0.442	0.927	5.530	0.000
Chen et al. (2016)	0.712	0.029	0.001	0.655	0.769	24.610	0.000
Cliff et al. (2007)	1.923	0.458	0.210	1.025	2.820	4.200	0.000
Cohen et al. (2015)	0.513	0.095	0.009	0.326	0.700	5.373	0.000
Gu et al. (2017)	1.040	0.116	0.013	0.813	1.266	8.997	0.000
Johnstone et al. (2017)	0.580	0.241	0.058	0.108	1.052	2.407	0.016
Lammle et al. (2016)	0.339	0.049	0.002	0.244	0.434	6.984	0.000
Laukkanen et al. (2015)	0.042	0.210	0.044	-0.370	0.454	0.202	0.840
Laurent et al. (2018)	1.562	0.430	0.185	0.720	2.405	3.635	0.000
Lee et al. (2020)a	1.570	0.416	0.173	0.754	2.386	3.770	0.000
Lee et al. (2020)b	0.909	0.293	0.086	0.335	1.483	3.104	0.002
Maskell et al. (2004)	0.143	0.304	0.092	-0.452	0.738	0.472	0.637
McKenzie et al. (1998)	0.647	0.091	0.008	0.468	0.825	7.107	0.000
Miller et al. (2015)	1.313	0.171	0.029	0.977	1.648	7.669	0.000
Miller et al. (2016)	0.869	0.201	0.040	0.475	1.263	4.325	0.000
Nathan et al. (2017)	0.789	0.157	0.025	0.482	1.097	5.029	0.000
Okely et al. (2017)	0.588	0.074	0.005	0.443	0.733	7.948	0.000
Platvoet et al. (2016)	0.427	0.106	0.011	0.219	0.634	4.030	0.000
Rudd et al. (2016)	0.480	0.118	0.014	0.250	0.710	4.083	0.000
Rudd et al. (2017)	0.437	0.189	0.036	0.067	0.808	2.312	0.021
Silveria et al. (2018)	0.442	0.335	0.112	-0.214	1.098	1.321	0.187
Skowroski et al. (2019)	1.485	0.411	0.169	0.679	2.291	3.610	0.000
	0.705	0.056	0.003	0.595	0.814	12.628	0.000

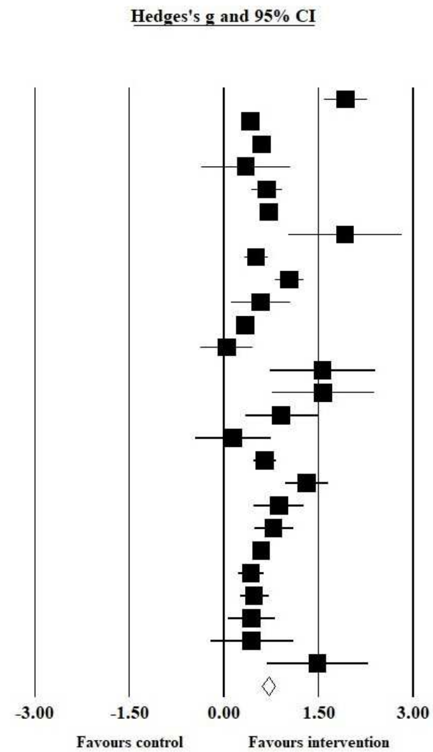

Figure 2.4 Forest plot for pooled effect sizes (Hedge's g) of all CSPAP-aligned PA intervention on children's MC from random effects meta-analysis of eligible studies.

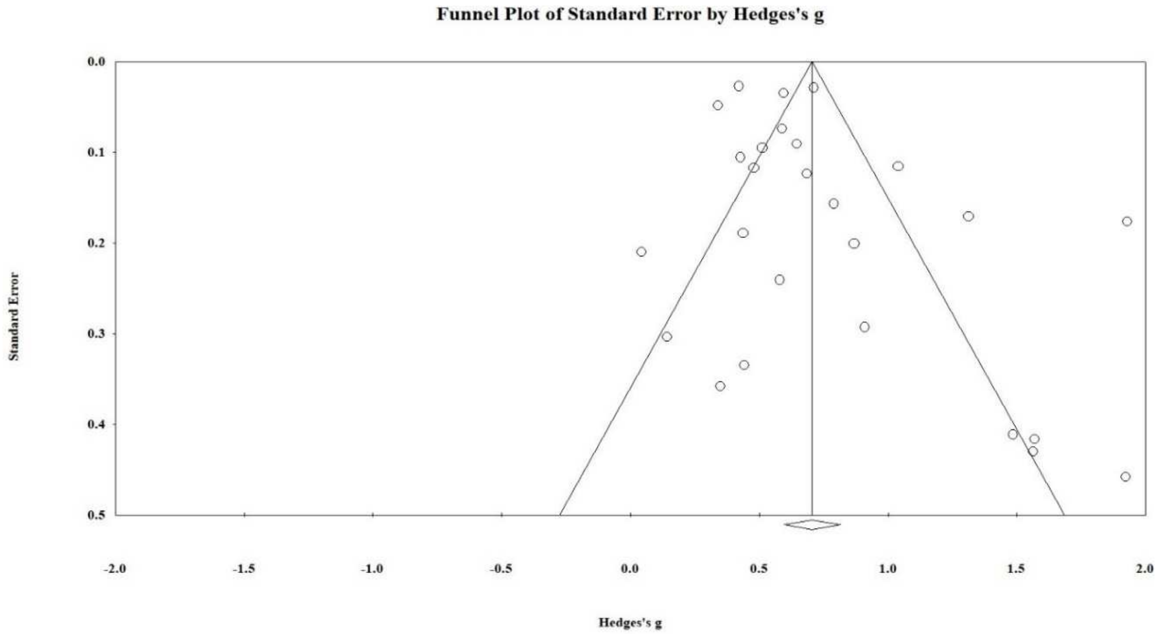

Figure 2.5 Funnel plot for assessment of publication bias using pooled analysis of effect sizes of MC outcomes.

Study name	Hedges's g	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
Bolger et al. (2019)	1.929	0.177	0.031	1.583	2.276	10.923	0.000
Burn et al. (2017)a	0.421	0.027	0.001	0.367	0.474	15.419	0.000
Burn et al. (2017)b	0.595	0.035	0.001	0.527	0.664	17.005	0.000
Burrows et al. (2014)	0.349	0.358	0.128	-0.353	1.051	0.975	0.329
Chan et al. (2019)	0.685	0.124	0.015	0.442	0.927	5.530	0.000
Cliff et al. (2007)	1.923	0.458	0.210	1.025	2.820	4.200	0.000
Cohen et al. (2015)	0.513	0.095	0.009	0.326	0.700	5.373	0.000
Johnstone et al. (2017)	0.580	0.241	0.058	0.108	1.052	2.407	0.016
Lee et al. (2020)a	1.570	0.416	0.173	0.754	2.386	3.770	0.000
Lee et al. (2020)b	0.909	0.293	0.086	0.335	1.483	3.104	0.002
Maskell et al. (2004)	0.143	0.304	0.092	-0.452	0.738	0.472	0.637
Miller et al. (2015)	1.313	0.171	0.029	0.977	1.648	7.669	0.000
Miller et al. (2016)	0.869	0.201	0.040	0.475	1.263	4.325	0.000
Nathan et al. (2017)	0.789	0.157	0.025	0.482	1.097	5.029	0.000
Rudd et al. (2016) TGMD	0.534	0.118	0.014	0.303	0.765	4.527	0.000
Rudd et al. (2017) TGMD	0.247	0.188	0.035	-0.121	0.615	1.317	0.188
Skowroski et al. (2019)	1.485	0.411	0.169	0.679	2.291	3.610	0.000
	0.787	0.082	0.007	0.626	0.948	9.570	0.000

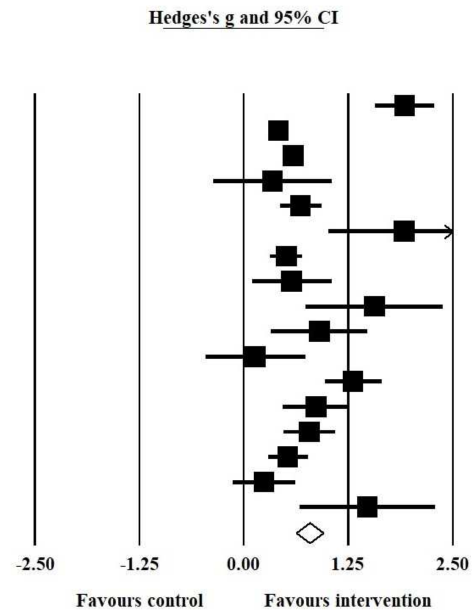

Figure 2.6 Forest plot for pooled effect sizes (Hedge’s g) of all CSPAP-aligned PA interventions using the TGMD-2 or -3 tool on children’s MC from random effects meta-analysis of eligible studies.

Study name	Statistics for each study						
	Hedges's g	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
Chen et al. (2016)	0.712	0.029	0.001	0.655	0.769	24.610	0.000
Gu et al. (2017)	1.040	0.116	0.013	0.813	1.266	8.997	0.000
Lammle et al. (2016)	0.339	0.049	0.002	0.244	0.434	6.984	0.000
Laukkanen et al. (2015)	0.042	0.210	0.044	-0.370	0.454	0.202	0.840
Laurent et al. (2018)	1.562	0.430	0.185	0.720	2.405	3.635	0.000
McKenzie et al. (1998)	0.647	0.091	0.008	0.468	0.825	7.107	0.000
Okely et al. (2017)	0.588	0.074	0.005	0.443	0.733	7.948	0.000
Platvoet et al. (2016)	0.427	0.106	0.011	0.219	0.634	4.030	0.000
Rudd et al. (2016) KTK	0.449	0.117	0.014	0.219	0.679	3.828	0.000
Rudd et al. (2017) KTK	0.510	0.190	0.036	0.138	0.882	2.685	0.007
Silveria et al. (2018)	0.442	0.335	0.112	-0.214	1.098	1.321	0.187
	0.572	0.077	0.006	0.421	0.722	7.448	0.000

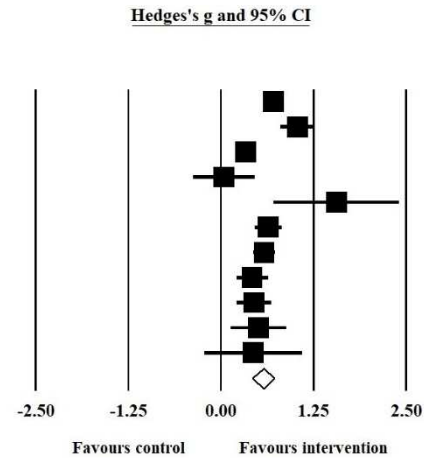

Figure 2.7 Forest plot for pooled effect sizes (Hedge's g) of all CSPAP-aligned PA interventions using other measurement tools on children's MC from random effects meta-analysis of eligible studies.

Study name	Hedges's g	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
	Statistics for each study						
Laukkanen et al. (2015)	0.042	0.210	0.044	-0.370	0.454	0.202	0.840
Platvoet et al. (2016)	0.427	0.106	0.011	0.219	0.634	4.030	0.000
Rudd et al. (2016)	0.449	0.117	0.014	0.219	0.679	3.828	0.000
Rudd et al. (2017)	0.510	0.190	0.036	0.138	0.882	2.685	0.007
Silveria et al. (2018)	0.442	0.335	0.112	-0.214	1.098	1.321	0.187
	0.406	0.067	0.005	0.274	0.538	6.034	0.000

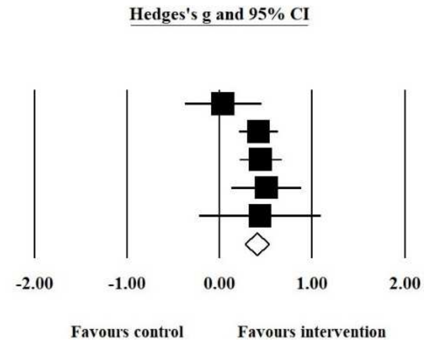

Figure 2.8 Forest plot for pooled effect sizes (Hedge’s g) of all CSPAP-aligned PA interventions using KTK tool on children’s MC from random effects meta-analysis of eligible studies.

Study name	Hedges's g	Upper limit	Standard error	Variance	Lower limit	Z-Value	p-Value
Bolger et al. (2019)	1.929	2.276	0.177	0.031	1.583	10.923	0.000
Chan et al. (2019)	0.685	0.927	0.124	0.015	0.442	5.530	0.000
Chen et al. (2016)	0.712	0.769	0.029	0.001	0.655	24.610	0.000
Gu et al. (2017)	1.040	1.266	0.116	0.013	0.813	8.997	0.000
Maskell et al. (2004)	0.143	0.738	0.304	0.092	-0.452	0.472	0.637
Miller et al. (2015)	1.313	1.648	0.171	0.029	0.977	7.669	0.000
Platvoet et al. (2016)	0.427	0.634	0.106	0.011	0.219	4.030	0.000
Rudd et al. (2016)	0.534	0.765	0.118	0.014	0.303	4.527	0.000
Rudd et al. (2017)	0.247	0.615	0.188	0.035	-0.121	1.317	0.188
	0.794	1.036	0.123	0.015	0.553	6.442	0.000

Statistics for each study

Hedges's g and 95% CI

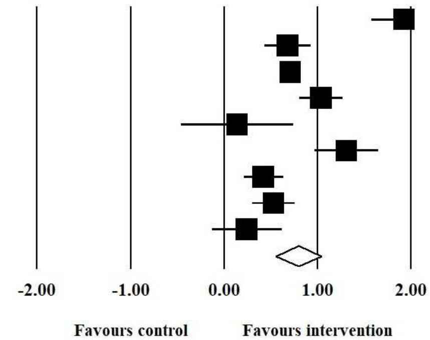

Figure 2.9 Forest plot for pooled effect sizes (Hedge's g) of PE single component interventions on children's total MC from random effects meta-analysis of eligible studies.

Study name	Statistics for each study						
	Hedges's g	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
Burrows et al. (2014)	0.349	0.358	0.128	-0.353	1.051	0.975	0.329
Johnstone et al. (2017)	0.603	0.241	0.058	0.130	1.075	2.498	0.012
Lammle et al. (2016)	0.339	0.049	0.002	0.244	0.434	6.984	0.000
Laukkanen et al. (2015)	0.042	0.210	0.044	-0.370	0.454	0.202	0.840
Laurent et al. (2018)	1.562	0.430	0.185	0.720	2.405	3.635	0.000
Lee et al. (2020)a	0.833	0.381	0.145	0.087	1.579	2.189	0.029
Lee et al. (2020)b	0.421	0.282	0.079	-0.131	0.973	1.495	0.135
Okely et al. (2017)	0.588	0.074	0.005	0.443	0.733	7.948	0.000
	0.483	0.101	0.010	0.286	0.681	4.796	0.000

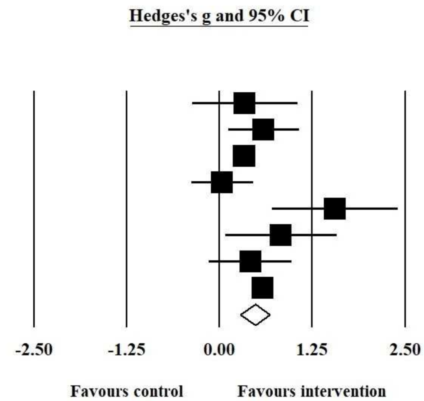

Figure 2.10 Forest plot for pooled effect sizes (Hedge’s g) of non-PE single component interventions on children’s total MC from random effects meta-analysis of eligible studies.

Study name	Statistics for each study							Hedges's g and 95% CI
	Hedges's g	Standard error	Variance	Lower limit	Z-Value	Upper limit	p-Value	
Lammle et al. (2016)	0.339	0.049	0.002	0.244	6.984	0.434	0.000	
Okely et al. (2017)	0.588	0.074	0.005	0.443	7.948	0.733	0.000	
Johnstone et al. (2017)	0.603	0.241	0.058	0.130	2.498	1.075	0.012	
	0.477	0.107	0.012	0.267	4.443	0.688	0.000	

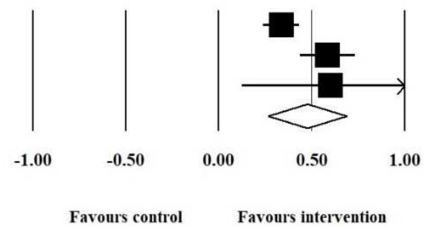

Figure 2.11 Forest plot for pooled effect sizes (Hedge's g) of PADS single component interventions on children's total MC from random effects meta-analysis of eligible studies.

Study name	Hedges's g	Standard error	Variance	Lower limit	Z-Value	Upper limit	p-Value
Burrows et al. (2014)	0.349	0.358	0.128	-0.353	0.975	1.051	0.329
Lee et al. (2020) Need	0.421	0.282	0.079	-0.131	1.495	0.973	0.135
Lee et al. (2020) FMS	0.833	0.381	0.145	0.087	2.189	1.579	0.029
	0.505	0.191	0.037	0.130	2.638	0.880	0.008

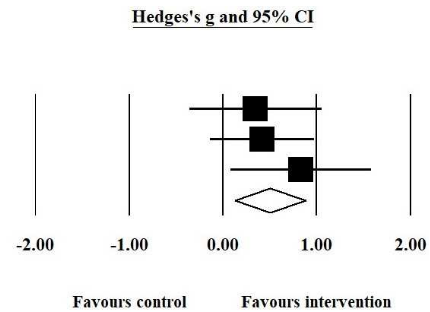

Figure 2.12 Forest plot for pooled effect sizes (Hedge's g) of PABAS single component interventions on children's total MC from random effects meta-analysis of eligible studies.

Study name	Hedges's g	Upper limit	Standard error	Variance	Lower limit	Z-Value	p-Value
			Statistics for each study				
McKenzie et al. (1998)	0.463	0.639	0.090	0.008	0.287	5.151	0.000
Miller et al. (2016)	0.869	1.263	0.201	0.040	0.475	4.325	0.000
Skowroski et al. (2019)	0.786	1.529	0.379	0.144	0.043	2.074	0.038
	0.634	0.936	0.154	0.024	0.332	4.115	0.000

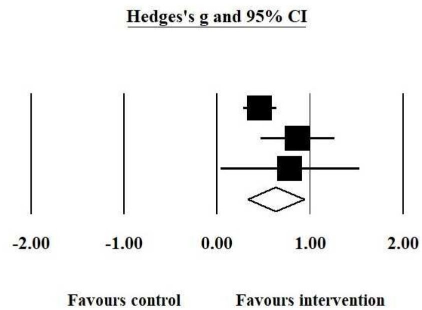

Figure 2.13 Forest plot for pooled effect sizes (Hedge's g) of PE and one additional CSPAP component (PE+1) interventions on children's total MC from random effects meta-analysis of eligible studies.

Study name	Hedges's g	Upper limit	Standard error	Variance	Lower limit	Z-Value	p-Value
Burn et al. (2017) Effect	0.421	0.474	0.027	0.001	0.367	15.419	0.000
Nathan et al. (2017)	0.789	1.097	0.157	0.025	0.482	5.029	0.000
Silveria et al. (2018)	0.442	1.098	0.335	0.112	-0.214	1.321	0.187
	0.543	0.816	0.139	0.019	0.271	3.908	0.000

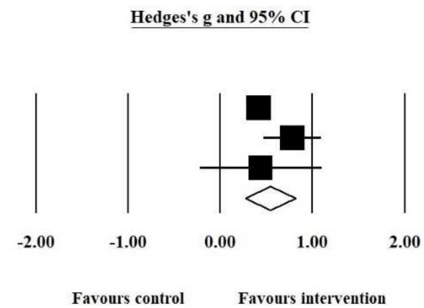

Figure 2.14 Forest plot for pooled effect sizes (Hedge’s g) of PE and two additional CSPAP component (PE+2) interventions on children’s total MC from random effects meta-analysis of eligible studies.

CHAPTER 3

STUDY 2: SYSTEMATIC REVIEW AND META-ANALYSIS OF PHYSICAL ACTIVITY INTERVENTIONS TO INCREASE ELEMENTARY CHILDREN'S SOCIAL AND EMOTIONAL LEARNING

INTRODUCTION

Physical activity (PA) is critical for the healthy development and well-being of school-aged youth (Physical Activity Guidelines Advisory Committee, 2018; Piercy et al., 2018; World Health Organization [WHO], 2018). Unfortunately, less than half (42.5%) of elementary-aged children (5-12 years) in the United States participate in the national guideline of accumulating at least 60 minutes of mostly moderate-to-vigorous PA each day (National Physical Activity Plan Alliance [NPAPA], 2018). Insufficient PA levels lead to limitations of necessary factors for children's health benefits (Pascoe et al., 2020). Further, it is clear that the importance of PA for children extends well beyond physical health and reaches into multiple dimensions of mental and psychosocial health (Ahn & Fedewa, 2011; Institute of Medicine, 2013; Lubans et al., 2016; Smedegaard et al., 2016). These dimensions include children's cognitive development, behavioral responses, and psychosocial functioning (Dale et al., 2019). The promotion of PA in children can represent great benefits that are carried throughout childhood, as it is a period when children are highly receptive to healthy behaviors and can continue these behaviors throughout their lifespan (Eime et al., 2013; Rodriguez-Ayllon et al., 2019).

Despite previous research on the connection between PA and physical and mental health (Andermo et al., 2020), it remains unclear what the most effective PA program approaches to sustain change are and how they can be successfully implemented to affect children's development.

Research on social emotional learning (SEL) might provide a path forward. Over the last decade, there has been increasing attention to the importance that SEL has on child development, with implications across learning, building and maintaining relationships, and early support for mental health and wellbeing (Currie, 2016; Wigelsworth et al., 2020; WHO, 2018). However, one in every six children (17.4%) in the U.S. has been diagnosed with a mental, behavioral, or developmental disorder (Cree et al., 2018), and these issues are more common among elementary school-aged children than older age groups (Ghandour et al., 2019). Such issues are problematic: they threaten the physical and psychological health of youth, diminish their ability to engage in learning and in society, and underscore the need to address SEL in school programs (Ghandour et al., 2019).

SEL has been used as an umbrella term. According to the Collaborative for Academic, Social, and Emotional Learning (CASEL), social and emotional development refers to "the process through all children and adults acquire and apply the knowledge, attitudes, and skills to develop healthy identities, manage emotions and achieve personal and collective goals, feel and show empathy for others, establish and maintain supportive relationships, and make responsible and caring decisions" (CASEL, 2020). In addition, Taylor et al. (2017) demonstrated that SEL competencies encompass a wide range of skills that help children recognize their feelings, establish healthy relationships, and solve

problems. There is consensus that schools are a central nexus through which SEL skills are developed and taught (Greenberg et al., 2017). A number of meta-analysis studies have shown that school-based SEL programs can improve children's positive behaviors, school engagement, and conduct problems while promoting desirable behaviors (Corcoran et al., 2018; Cefai et al., 2018; Durlak et al., 2011; Greenberg et al., 2017; Lewis et al., 2021; Wigelsworth et al., 2016).

The school environment serves as a strategic point of intervention to provide children with PA opportunities and SEL (Greenberg et al., 2017; Kunth et al., 2012; Rasberry et al., 2011). However, only 15% of elementary schools require or recommend that students take PE a minimum of three times per week for the whole school year; additionally, more than one third (35.2%) of school districts have no policy requiring the provision of regularly scheduled recess/breaks in elementary schools (NPAPA, 2018). Thus, other contexts beyond PE lessons are needed to leverage the impact of PE on children's PA and SEL. A promoted strategy to increase daily PA is the implementation of a comprehensive school PA program (CSPAP). A CSPAP is recommended as a national framework to help all children be physically active at least 60 minutes each day as well as develop their targeted PE outcomes (CDC, 2019). CSPAPs are single- or multicomponent approaches to PA promotion that target settings and resources based in, and connected with, schools, families, and communities (CDC, 2019).

As noted by Jones et al. (2018), SEL occurs more often when it is intentionally planned for and when teaching directly reflects the desired outcomes. Previous reviews indicate the promising effects of PA interventions on children's SEL outcomes (e.g., psychological variables), highlighting purposeful intervention design using multiple

strategies, such as enjoyment and task-oriented teaching style (Barry, Clarke, & Dowling, 2017; Burns, Fu, & Podlog, 2017; Klos et al., 2020). However, there is a lack of understanding of effective CSPAP-aligned PA program approaches to improve SEL in children. In this sense, it is crucial to consider the aspects of different types of CSPAP-aligned interventions to better understand what may or may not work to improve SEL. Presenting an up-to-date review may allow a better understanding of the effects of these interventions on children's SEL. The purpose of this study was to conduct a systematic review and meta-analysis of the effectiveness of PA interventions in increasing the SEL of elementary school children (5-12 years). The CSPAP framework was used to consider the effectiveness of different intervention approaches on children's SEL. This review will better elucidate connections between SEL and different facets of CSPAP programming and will allow such conceptualizations of CSPAPs to evolve toward reflecting the many ways students' SEL can be supported. An additional benefit of this review will be to articulate different types of PA interventions that produce positive changes in children's SEL, which may help bridge the gap between theory and practice for CSPAP practitioners.

METHOD

Registration and protocol

This study followed the updated Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Appendix D; Page et al., 2021) with additional recommendations for systematic meta-reviews (Hennessy et al., 2019) and was registered with the International Prospective Register of Systematic Reviews at https://www.crd.york.ac.uk/prospero/ (registration number CRD42021224998).

Inclusion/eligibility criteria

Studies with the following characteristics were included in our review:

1. Participants were aged 5-12 years (primary/elementary school);
2. PA interventions primarily focused on improving and assessing SEL components;
3. Type of interventions: Any school-, home-, or community-based interventions for children with clear intent to improve SEL proficiency;
4. Type of studies: Employed a Randomized Controlled Trials (RCTs) design, Cluster-RCTs (C-RCTs) or rigorous (matched or statistically controlled) quasi-experimental design (i.e., N-RCT).

Exclusion criteria

Studies with the following characteristics were excluded from the review:

1. Studies that reported on populations of focus that were outside the age range defined above; or participants were not 'typically developing' (i.e. had a clinically diagnosed physical or intellectual disability or condition affecting movement, e.g. cerebral palsy, traumatic brain injury/concussion);
2. Studies that did not aim to improve and assess at least one SEL outcome were excluded;
3. Studies reported as abstracts, theses/dissertations and unpublished literature were excluded.

Search strategy and terms

The studies were obtained in February 28, 2021 using six electronic databases: PubMed/Medline, Embase, ERIC, SPORTDiscus, CINAHL, and PsycINFO. The search

strategy consisted of four elements: study population (e.g., elementary school student), study design (e.g., RCT and quasi-experimental design), intervention (e.g., PA and exercise), and outcome measures (Appendix E). The search was limited to peer-reviewed academic journal articles published in English in all available years.

Data extraction/collection process

Data were imported into Endnote X9.3 and duplicates were removed. The selected references were imported to a web□based software platform that streamlines the production of systematic reviews (Covidence systematic review software, Veritas Health Innovation, Melbourne, Australia, available at www.covidence.org). Two independent reviewers screened the titles and abstracts of retrieved records for possible inclusion. Of the records identified as possibly eligible, the full texts were obtained, and two independent reviewers assessed the records for inclusion. For each included study, two reviewers extracted data into a Microsoft Excel (Microsoft Corporation, Redmond, WA, USA) data collection form. Data were extracted on the following: the study design; the intervention approach, based on the CSPAP framework; the sample size; the intervention characteristics including session duration, frequency, length, delivery, and the name of programs; the types and methods of measured outcomes; the fidelity of implementation measure; and the main results in SEL outcomes (Appendix F). For all steps in the screening process and data extraction, a third reviewer checked the data for errors, and discrepancies were resolved through discussion and consensus of judgement. If data were missing, authors were contacted.

Qualitative data synthesis

Extracted results showed information on the article reference, study design, intervention approach (i.e., type of CSPAP components used), study characteristics (country, school setting, school level, participants, deliverer, SEL outcomes, etc.), main results, and fidelity report. Results were organized into three sections by: (1) study design (i.e., C-RCT/RCT and N-RCT), (2) single CSPAP component interventions (e.g., PE, PADS, etc.) and (3) the number and combination of CSPAP components used in multicomponent interventions (e.g., PE and PADS). Furthermore, we used the Collaborative for Academic Social and Emotional Learning (CASEL, 2020) as an analytical framework for categorizing the different SEL outcomes from included studies. There are five sets of interrelated social and emotional competencies, each comprised of a range of skills, abilities, and knowledge (Osher et al., 2016): self-awareness (SA), self-management (SM), social-awareness (SoA), relationship skills (RS), and responsible decision making (RDM). Using this framework, two reviewers categorized each SEL outcome from the studies into each of the five competencies.

Quantitative data synthesis

Effect sizes were calculated for the intervention group relative to the comparison group for each study. When the necessary data were not available in the original article, we requested it from the authors. If data could still not be obtained, we extracted the data from the graphs when available. If that was not possible, we excluded the study from the quantitative analysis. A meta-analysis for a given SEL outcome was conducted if at least three studies reported interventions addressing the same CSPAP components and

provided sufficient data for the calculation of effect size. Reported SEL outcomes were reorganized by the five competencies in the CASEL framework.

Pre- and post-intervention mean ± standard deviation (SD) for a given SEL outcome, and sample size from each study were converted to Hedges' g effect size (Hedges, 1981). Specifically, we calculated standardized mean differences both for outcome scores at the end of the intervention period (i.e., post-intervention) and change-from-baseline (pre-intervention) outcomes. Scores post-intervention effect sizes refer to intervention group results compared with comparison or control group results after interventions. We did not include follow-up assessment data. In all analyses, we used the random-effects model to account for differences between studies that might impact the treatment effect (Deeks et al., 2008; Kontopantelis et al., 2013). The effect size values are presented alongside their respective 95% Confidence Intervals (CIs). Calculated effect sizes were interpreted using the following scale: small ($g<.40$), moderate ($g = .40–.07$), and large ($g>.07$), according to the Cochrane Handbook (Higgins et al., 2019). Heterogeneity (i.e., between studies variability) was evaluated using the I-squared (I^2) statistic. I^2 values of <25%, 25-75%, and > 75% were considered to represent low, moderate, and high levels of heterogeneity, respectively (Huedo-Medina, Sanchez-Meca, Marin-Martinez, & Botella, 2006). The risk of bias was explored using the visual inspection of funnel plots and Egger's regression test (Egger, Smith, Schneider, & Minder, 1997). Publication bias was not produced as the meta-analyses included <10 studies/interventions (Higgins & Green, 2017).

A series of models were analyzed to address the following: (1) the pooled effect of PA interventions across all studies on elementary school-age children's SEL (overall

and by competency), (2) the pooled effect of single CSPAP component interventions (PE and PABAS, respectively) on children's SEL (overall and by competency), and (3) the pooled effect of all single component interventions compared to the pooled effect of all multicomponent interventions on children's SEL (overall). In addition, moderation analyses were performed to explore the impact of potential explanatory variables and moderators (intervention duration [<6 months vs. ≥ 6 months], delivery agent [research team vs. school-based team vs. combined], and study design [C-RCT/RCT vs. N-RCT]) on the effect sizes with meta-regressions when sufficient data were available (i.e., at least ten studies for each explanatory variable; Higgins et al., 2019). The results were expressed as regression coefficients (β) estimates, 95% CIs, and the p value. All analyses were carried out using the Comprehensive Meta-Analysis program (version 3.3.070; Biostat, Englewood, NJ, USA). The statistical significance threshold was set at $p<.05$.

Risk of bias assessment

Risk of bias in the included studies was assessed by two reviewers independently through discussion using the Cochrane Risk of Bias Tool (RoB 2.0) with additional considerations for C-RCTs and RCTs (Sterne et al., 2019), which consists of five domains and an overall judgment. The five domains are: 1) bias arising from the randomization process; 2) bias due to deviations from the intended interventions; 3) bias due to missing outcome data; 4) bias in measurement of the outcome; and 5) bias in selection of the reported result. Based on the answers (yes, probably yes, probably no, no, not applicable, no information) to a series of signaling questions in the guidance document, the judgment options within each domain consist of "low risk of bias," "some concerns" or "high risk of bias" (Sterne et al., 2019).

The N-RCTs (i.e., quasi-experimental) studies were assessed with the Risk of Bias in Non-randomized Studies of Interventions tool (ROBINS-□; Sterne et al., 2016), which consists of seven domains and an overall judgement. The seven domains are: 1) bias due to confounding; 2) bias in selection of participants into the study; 3) bias in classification of interventions; 4) bias due to deviations from intended intervention; 5) bias due to missing data; 6) bias in measurement of outcomes and 7) bias in selection of the reported result (Sterne et al., 2016). Domain specific risk of bias assessment was used to judge the overall risk of bias for each study. Disagreements between reviewers were resolved through discussion and consensus by a third evaluator. Before correcting for observed differences, the agreement between reviewers was assessed using a Kappa correlation for risk of bias ($k > 0.8$). A risk of bias graph was made via the *robvis* R package (McGuinness & Higgins, 2021).

RESULTS

A total of 7,783 search records were initially identified. The authors screened 5,092 records after removing duplicate records. This first level of screening, separated by title and abstract, identified 5,051 full-text articles to be reviewed for eligibility. Ultimately, of the remaining 78 articles, 32 studies were included in the qualitative synthesis and 27 studies were included in quantitative synthesis. The process of literature identification and selection is outlined in the PRISMA flowchart (Figure 3.1; Page et al., 2021). The quality assessment for C-RCTs or RCTs revealed 10 studies as low risk in quality, 8 studies as having some concerns in quality, and no study as high risk (Figure 3.2). For N-RCTs, 12 studies were evaluated as low risk of bias, 3 studies as moderate quality, and no study as serious risk of bias (Figure 3.3). Overall, the problems were

related to confounders, either in the participant selection process or inequality among participants, and authors did not report how they handled missing outcome data.

Characteristics of the included studies

Across the 32 studies, there were a total of 7,791 participants (45% female, 55% male) from 152 classes and 193 schools. The sample size ranged from 29 (Olive et al., 2020) to 925 participants (Grillich et al., 2016) with the age of intervention children ranging from 5 to 12 years. 11 studies were conducted in North America (i.e., Mexico and United States), 13 in Europe (i.e., Belgium, Denmark, England, Greece, Italy, Serbia, and Spain), 3 in Asia (i.e., Hongkong and Singapore), 2 in Australia, and 1 in South America (i.e., Brazil). Additionally, 12 studies were conducted in urban settings, 2 in rural and urban settings, and 1 in a rural setting. The setting was not specified in 17 of the studies. For the measurement of SEL-related outcomes, 24 studies measured SA, 14 studies measured SM, 14 studies measured SoA, 13 studies measured RS, and 6 studies measured RDM (CASEL, 2020).

Descriptions of CSPAP components

Study design. 10 C-RCTs (31%), 7 RCTs (22%), and 15 N-RCT studies (47%) were included in this review.

C-RCT/RCT. 17 studies were C-RCTs or RCTs, with the average number of schools and classes across all studies being 10 (range 1–45) and 15 (range 4–53), respectively (Andrade et al., 2020; Duncan et al., 2009; Grillich et al., 2016; Mayorga-Vega et al., 2012). The average sample size was 283 students (range 68–925; Duncan et al., 2009; Grillich et al., 2016), and the total number of students was 4,804 (50% Female, 50% Male). The average intervention duration, frequency, and length was 16 weeks

(range 2–48; Andrade et al., 2020; Grillich et al., 2016), 3 times per week (range 1–5; Duncan et al., 2009; Elbe et al., 2017; Ho et al., 2017; Huang et al., 2018; Lwin & Malik, 2014), and 55 minutes per session (range 40–90; Andrade et al., 2020; Duncan et al., 2009, 2018; Ho et al., 2017), respectively.

N-RCT. 15 studies were N-RCTs, with the average number of schools and classes across all studies being 5 (range 1–21; Brown & Fry, 2014; Duck et al., 2020; Fu & Burns, 2018); Shachar et al., 2016) and 5 (range 1–12; Escartí et al., 2010; Pesce et al., 2013), respectively. The average sample size was 199 students (range 29–877; Olive et al., 2020; Wong et al., 2016), and the total number of students was 2,987 (38% Female, 62% Male). The average intervention duration, frequency, and length was 18 weeks (range 4–36; Escartí et al., 2010; Griffiths & Griffiths, 2019; Olive et al., 2020), 2 times per week (range 1–5; Brown & Fry, 2014; Griffiths & Griffiths, 2019; Hassandra & Goudas (2010); Pesce et al., 2013; Shachar et al., 2016), and 61 minutes per session (range 30–120; Brown & Fry, 2014; Fu & Burns, 2018), respectively.

Single CSPAP component interventions.

Considering the CSPAP framework, 15 studies out of 28 (54%) used only PE for the intervention (Andrade et al., 2020; Buišić & Đorđić, 2018; Chan et al., 2019; Cocca et al., 2020; Duncan et al., 2018; Escartí et al., 2010; Goudas & Magotsiou, 2009; Griffiths & Griffiths, 2019; Hassandra & Goudas, 2010; Lwin & Malik, 2014; Mayorga-Vega et al., 2012; Moreno-Murcia & Sánchez Latorre, 2016; Olive et al., 2019; PÉRez-OrdÁS et al., 2020; Pesce et al., 2013; Duncan et al., 2009; Ho et al., 2017; Huang et al., 2018; Olive et al., 2020; Shachar et al., 2016; Wong et al., 2016), 10 studies (36%) used only PABAS (Annesi et al., 2016a, 2016b, 2017; Brown & Fry, 2014; Duncan et al., 2009; Ho

et al., 2017; Huang et al., 2018; Olive et al., 2020; Shachar et al., 2016; Wong et al., 2016), and 3 studies (10%) used only PADS (Fu & Burns, 2018; Grillich et al., 2016; Mendelson et al., 2010). No study used only SI or FCE component for the intervention.

PE. For the studies that used only PE as the intervention approach, the average number of schools and classes across all studies were 5 (range 1–8; Andrade et al., 2020; Mayorga-Vega et al., 2012; Pesce et al., 2013) and 7 (range 1–12; Escartí et al., 2010; Lwin & Malik, 2014; Pesce et al., 2013), respectively. The average sample size was 227 students (range 41–821; Hassandra & Goudas, 2010; Olive et al., 2020), and the total number of students was 3,406 (47% Female, 53% Male). The average intervention duration, frequency and length was 16 weeks (range 2–36; Andrade et al., 2020; Escartí et al., 2010), 2 times per week (range 1–3) and 51 minutes per session (range 40–70; Andrade et al., 2020; Chan et al., 2019; Duncan et al., 2018), respectively.

PE interventions involved a program that integrated movement-based video games that aimed to promote children's psychological health during PE lessons (Andrade et al., 2020; Lwin & Malik, 2014); Teaching Personal and social Responsibility (TPSR) models in PE lessons that used group discussion and self-reflection (Buišić & Đorđić, 2018) and field hockey lessons (Hassandra & Goudas, 2010) to develop students' personal and social responsibility (e.g., self-control, effort, problem-solving, and self-regulation, etc.); a study that investigated the efficacy of a 5 level progression approach for a TPSR model in PE lessons (e.g., football, volleyball, and badminton unites) using different strategies (e.g., group meeting, reflection time, etc.; Escartí et al., 2010; PÉRez-OrdÁS et al., 2020); an assessment-based intervention program that focused on fun movement activities, mastery, and support to increase students' perceptions of physical

and movement skill competence, teacher support, and enjoyment (Chan et al., 2019); game-based PE lessons that emphasized team or individual challenges, modified ball games, cooperation, and skill games, as well as inquiry games to provide stimulating challenges to students (Cocca et al., 2020); a program that consisted of mobility-focused warm-up exercises (deep squat and bear crawls), followed by a series of fundamental movement exercises (Duncan et al., 2018); a cooperative PE program to develop children's social skills and attitudes toward group work (Goudas & Magotsiou, 2009); a multiphase intervention that used video sport role models, circuit training sessions within PE lessons, and PA diaries that recorded activity at school, at home, and in the community (Griffiths & Griffiths, 2019); a program focused on physical self-concept and physical fitness elements (Mayorga-Vega et al., 2012); approaches that used autonomy support strategies (i.e., providing challenges and options) based on self-determination theory (Moreno-Murcia & Sánchez Latorre, 2016); guided discovery methods that challenged children with movement tasks and emphasized balance and fundamental motor skills (Olive et al., 2019); and PE lessons including multiple sports taught through four didactic modules (pre-tumbling, rhythmic gymnastics, ball mini-games, and dexterity circuits). Across all studies, SEL outcomes included SA (n=12), SM (n=8), SoA (n=9), RS (n=6) and RDM (n=3).

The intervention deliverer varied across interventions. 3 studies (20%) were delivered by a research team, 11 studies (73%) by a school-based team (i.e., PE teachers, classroom teachers, students), and 1 study (7%) by a combined team (i.e., fitness expert, research team, and teachers). Additionally, 4 studies (27%) reported fidelity of intervention using observation, documentation, weekly meetings, and evaluation (Chan et

al., 2019; Duncan et al., 2018; Moreno-Murcia & Sánchez Latorre, 2016; PÉRez-OrdÁS et al., 2020).

PABAS. For the studies that used only PABAS as the intervention approach, the average number of schools and classes across all studies was 7 (range 1–21; Brown & Fry, 2014; Duncan et al., 2009; Shachar et al., 2016). The average sample size was 247 students (range 29–877; Olive et al., 2020; Wong et al., 2016), and the total number of students was 2,470 (40% Female, 60% Male). The average intervention duration, frequency, and length was 14 weeks (range 4–36; Annesi et al., 2016a; Olive et al., 2020), 3 times per week (range 1–5; Brown & Fry, 2014; Ho et al., 2017; Huang et al., 2018) and 66 minutes per session (range 40–120; Brown & Fry, 2014; Duncan et al., 2009), respectively.

PABAS interventions involved an after school program based on a social cognitive theoretical approach that included physical activities and developed self□management/self□regulatory skills via self□talk, establishing social support, and goal setting to foster self-efficacy (Annesi et al., 2016a, 2016b, 2017); an after school program designed to encourage interaction, team work, and giving best effort and discussions/group activities designed to cultivate positive life skills rather than personal performance (Brown & Fry, 2014); an after school program based on positive youth development theories that focused on students’ sports mentorship and establishing a constructive environment (e.g., supportive relationships and positive social norms; Ho et al., 2017); the incorporation of a range of plyometric type training/exercises as extra after school club activity (Duncan et al., 2009); an after-school program that utilized Bandura’s social cognitive theory and active commuting designed to enhance children’s

self-efficacy, parent child self-efficacy, parent self-efficacy, and parent outcome expectations (Huang et al., 2018); physical activity and games (PAGs) lessons that followed a TPSR model infused with social–emotional and character development activities (e.g., awareness talk, group meeting) during after school programs to enhance students' self-perceptions (Olive et al., 2020); after school activities that included martial arts and group sport activities (e.g., soccer, basketball, volleyball) to enhance students' self-control skills (Shachar et al., 2016); and an after school program included different physical activities performed in small groups to improve students' agility, balance, and core strength (Wong et al., 2016). Across the studies, SEL outcomes included SA (n=8), SM (n=4), SoA (n=2), RS (n=2), and RDM (n=2).

Only 1 study (10%) was delivered by a research team (i.e., graduate students and research assistants), another study (10%) was delivered by a school-based team (i.e., PE teachers, classroom teachers, and students), and 8 studies (80%) by a combined team (i.e., after□school care counsellors, certified sports trainers, exercise physiologist, research staff, and parents). Additionally, 5 studies (50%) reported fidelity of intervention using a survey, observation, and/or fidelity checklist and rubric (Annesi et al., 2016a, 2016b, 2017; Ho et al., 2017; Olive et al., 2020).

PADS. For the studies that used only PADS as the intervention approach, the average number of schools and classes across all studies were 17 (range 1–45) and 28 (range 2–53), respectively (Fu & Burns, 2018; Grillich et al., 2016). The average sample size was 359 students (range 66–925; Fu & Burns, 2018; Grillich et al., 2016), and the total number of students was 1,078 (52% Female, 48% Male). The average intervention duration, frequency and length was 26 weeks (range 12–48; Grillich et al., 2016;

Mendelson et al., 2010), 3 times per week and 38 minutes per session (range 30–45; Fu & Burns, 2018; Mendelson et al., 2010), respectively.

The PADS interventions involved curriculum that utilized both social cognitive theory and Active Video Gaming (AVG) lessons (i.e., Adventure to Fitness, Cosmic Kids Yoga, GoNoodle, Dance, Dance, Revolution, and Wii Fit) that were incorporated as part of students' regular school class time (Fu & Burns, 2018); a program that implemented the World Health Organization's Health Promoting School framework program aimed to create a healthy and positive learning environment by classroom teachers (Grillich et al., 2016); and a resource program occurring during the school day that included yoga-based PA, breathing techniques, and guided mindfulness activities designed to enhance children's capacities for sustained attention, promote greater awareness, and increased regulation of cognition (Mendelson et al., 2010). SEL outcomes included SA (n=2), SM (n=2), SoA (n=2), RS (n=2), and RDM (n=1).

2 studies (66%) were delivered by a research team (Fu & Burns, 2018; Mendelson et al., 2010) and 1 study (34%) by a school-based team (i.e., classroom teachers; Grillich et al., 2016). Additionally, only 1 study (33%) reported fidelity of intervention using a survey (Grillich et al., 2016).

Multiple CSPAP components interventions.

A total of 4 studies (13%) used intervention approaches that could be mapped onto multiple components within the CSPAP framework (Bundy et al., 2017; Coolkens et al., 2018; Duck et al., 2020; Elbe et al., 2017). The most commonly used components in multicomponent interventions were PADS (n=3) followed by PE (n=2). PABAS (n=1)

and FCE (n=1) were included in less than half of the multicomponent studies. No study included three or more CSPAP components.

PE + 1 additional CSPAP component. 2 of the studies (50%) reported an intervention that included PE and 1 additional CSPAP component. 1 study included PE and PABAS (Elbe et al., 2017) and 1 study included PE and PADS (Coolkens et al., 2018). The average number of schools and classes across all studies were 11 and 14, respectively. The average sample size was 290 students and the total number of students was 581 (52% Female, 48% Male). The average intervention duration, frequency, and length was 30 weeks, 2 times per week and 80 minutes per session, respectively (Coolkens et al., 2018; Elbe et al., 2017).

PE+1 interventions involved a program in which PE teachers provided parkour-related games during PE lessons and recess time (Coolkens et al., 2018) and a program that used team sports (e.g., small-sided football and basketball) and individual sport sessions (e.g., strength training or interval running) during PE and recess time or after school (Elbe et al., 2017). SEL outcomes included SA (n=1), SoA (n=1), and RS (n=1).

1 study (50%) was delivered by a school-based team (i.e., PE teachers; Coolkens et al., 2018) and 1 study (50%) by a combined team (i.e., university staff members and teachers; Elbe et al., 2017). Additionally, 1 study (33%) reported fidelity of intervention using observation (Coolkens et al., 2018).

Multicomponent interventions without PE. 2 of the studies (50%) reported a multicomponent intervention with no PE component. 1 study included the combination of PADS and PABAS (Bundy et al., 2017) and 1 study included the combination of PADS and FCE (Duck et al., 2020). The average number of schools and classes across all

studies were 7 and 7, respectively. The average sample size was 128 students and the total number of students was 256 (49% Female, 51% Male). The average intervention duration was 12 weeks. Neither study reported frequency or length of intervention (Bundy et al., 2017; Duck et al., 2020).

The intervention programs consisted of placing recycled materials without an obvious play purpose on school playgrounds (Bundy et al., 2017) and a teacher□led experience that utilized regular classroom time and home activities that leveraged easy□to□use technology and the standards□based UNICEF Power Up curriculum using active games, videos, and yoga (Duck et al., 2020); SEL outcomes included SA (n=1) and RS (n=2). Both interventions were delivered by a school-based team (i.e., teachers and students). Neither study reported fidelity of intervention.

Meta-Analysis

Effectiveness across all interventions on SEL (overall and by competence)

Total SEL. The meta-analysis for 26 studies indicated a statistically significant and moderate pooled intervention effect on children's total SEL (Hedge's g = 0.44, 95% CI=0.34–0.54; p=.000; I^2 = 68.5 %; Figure 3.4). The relative weight of each study in the analysis ranged from 1.99 to 6.59 %. For all included studies, Egger's regression test for asymmetry of the funnel plot was not significant (β = 0.87, p=0.17), indicating no evidence of publication bias (Figure 3.5). Results from the meta-regression found that intervention duration (β = -0.17; 95% CI=-0.37–0.03; p=0.09), delivery agent (β = -0.11; 95% CI=-0.48–0.26; p=0.30), and study design (β = -0.01; 95% CI=-0.21–0.20; p=0.69) were not statistically significant moderator variables/factors affecting overall study effect sizes (i.e., children's total SEL).

In a subsequent analysis, the studies were analyzed by SEL competency.

SA. The meta-analysis for 21 studies indicated a statistically significant and moderate pooled intervention effect on children's SA (Hedge's g = 0.43, 95% CI=0.34–0.52; p=.000; I^2 = 55.7 %; Figure 3.6). The relative weight of each study in the analysis ranged from 1.75 to 8.04 %. Results from the meta-regression found that intervention duration (β = -0.20; 95% CI=0.17–0.43; p=0.11), delivery agent (β = 0.22; 95% CI=0.23–0.54; p=0.32), and study design (β = -0.01; 95% CI=0.11–0.84; p=0.49) were not found to be a statistically significant moderator variables/factors affecting overall study effect sizes (i.e., children's SA).

SM. The meta-analysis for 12 studies indicated a statistically significant and small pooled intervention effect on children's SM (Hedge's g = 0.36, 95% CI=0.22–0.50; p=.000; I^2 = 71.6 %; Figure 3.7). The relative weight of each study in the analysis ranged from 5.07 to 13.46 %. Results from the meta-regression found that intervention duration (β = -0.17; 95% CI=0.24–0.76; p=0.27), delivery agent (β = -0.15; 95% CI=0.27–0.64; p=0.31), and study design (β = -0.01; 95% CI=-0.12–0.22; p=0.09) were not found to be a statistically significant moderator variables/factors affecting overall study effect sizes (i.e., children's SM).

SoA. The meta-analysis for 9 studies indicated a statistically significant and moderate pooled intervention effect on children's SoA (Hedge's g = 0.44, 95% CI=0.30–0.59; p=.001; I^2 = 72.3 %; Figure 3.8). The relative weight of each study in the analysis ranged from 4.36 to 17.21 %.

RS. The meta-analysis for 11 studies indicated a statistically significant and moderate pooled intervention effect on children's RS (Hedge's g = 0.44, 95% CI=0.29–

0.58; p=.001; I^2 = 72.7 %; Figure 3.9). The relative weight of each study in the analysis ranged from 5.07 to 14.49 %. Results from the meta-regression found that intervention duration (β = -0.12; 95% CI=0.14–0.59; p=0.24), delivery agent (β = -0.08; 95% CI=0.33–0.65; p=0.25), and study design (β = -0.05; 95% CI=-0.07–0.32; p=0.19) were not found to be a statistically significant moderator variables/factors affecting overall study effect sizes (i.e., children's RS).

RDM. The meta-analysis for 5 studies indicated a statistically significant and moderate pooled intervention effect on children's RDM (Hedge's g = 0.64, 95% CI=0.52–0.76; p=.000; I^2 = 0.0 %; Figure 3.10). The relative weight of each study in the analysis ranged from 4.10 to 58.51 %.

Single CSPAP component intervention on SEL (overall and by competence)

Total SEL. 14 studies used only PE as the intervention approach to address any SEL outcome. The meta-analysis for studies indicated a statistically significant and large pooled intervention effect on children's total SEL (Hedge's g = 0.41, 95% CI=0.29–0.53; p=.000; I^2 = 48.7%; Figure 3.11). The relative weight of each study in the analysis ranged from 5.01 to 9.36%. Results from the meta-regression found that intervention duration (β = -0.17; 95% CI=-0.37–0.03; p=0.09), delivery agent (β = -0.11; 95% CI=-0.48–0.26; p=0.30), and study design (β = -0.01; 95% CI=-0.21–0.20; p=0.69) were not found to be a statistically significant moderator variables/factors affecting overall study effect sizes (i.e., children's total SEL).

In a subsequent analysis, the studies that used only PE as the intervention approach were analyzed by competency.

SA. 12 studies used only PE as the intervention approach to address SA. The meta-analysis for studies indicated a statistically significant and moderate pooled intervention effect on children's SA (Hedge's g = 0.43, 95% CI=0.32–0.54; p=.000; I^2 = 42.1%; Figure 3.12). The relative weight of each study in the analysis ranged from 3.17 to 16.64%. Results from the meta-regression found that intervention duration (β = -0.15; 95% CI=0.18–0.69; p=0.25), delivery agent (β = -0.25; 95% CI=-0.12–0.15; p=0.10), and study design (β = -0.07; 95% CI=-0.10–0.42; p=0.29) were not found to be a statistically significant moderator variables/factors affecting overall study effect sizes (i.e., children's RS).

SM. 8 studies used only PE as the intervention approach to address SM. The meta-analysis for studies indicated a statistically significant and moderate pooled intervention effect on children's SM (Hedge's g = 0.41, 95% CI=0.24–0.57; p=.000; I^2 = 67.7%; Figure 3.13). The relative weight of each study in the analysis ranged from 5.53 to 16.81%.

SoA. 8 studies used only PE as the intervention approach to address SoA. The meta-analysis for studies indicated a statistically significant and moderate pooled intervention effect on children's SoA (Hedge's g = 0.38, 95% CI=0.27–0.49; p=.000; I^2 = 40.3%; Figure 3.14). The relative weight of each study in the analysis ranged from 3.06 to 29.77%.

RS. 5 studies used only PE as the intervention approach to address RS. The meta-analysis for studies indicated a statistically significant and large pooled intervention effect on children's RS (Hedge's g = 0.45, 95% CI=0.23–0.67; p=.000; I^2 = 76.4%;

Figure 3.15). The relative weight of each study in the analysis ranged from 14.33 to 28.27%.

RDM. 3 studies used only PE as the intervention approach to address RDM. The meta-analysis for studies indicated a statistically significant and large pooled intervention effect on children's RMD (Hedge's g = 0.55, 95% CI=0.31–0.78; p=.000; I^2 = 0.0%; Figure 3.16). The relative weight of each study in the analysis ranged from 15.97 to 47.93%.

Total SEL 8 studies used only PABAS as the intervention approach to address any SEL outcome. The meta-analysis for studies indicated a statistically significant and large pooled intervention effect on children's total SEL (Hedge's g = 0.52, 95% CI=0.31–0.73; p=.000; I^2 = 75.5%; Figure 3.17). The relative weight of each study in the analysis ranged from 5.75 to 17.11%.

In a subsequent analysis, the studies that used only PABAS as the intervention approach were analyzed by competency.

SA. 7 studies used only PABAS as the intervention approach to address SA. The meta-analysis for studies indicated a statistically significant and moderate pooled intervention effect on children's SA (Hedge's g = 0.49, 95% CI=0.36–0.62; p=.000; I^2 = 30.9%; Figure 3.18). The relative weight of each study in the analysis ranged from 6.14 to 24.35%.

SM. 3 studies used only PABAS as the intervention approach to address SM. The meta-analysis for studies indicated a statistically significant and small pooled intervention effect on children's SM (Hedge's g = 0.25, 95% CI=0.12–0.38; p=.000; I^2 = 0.0%; Figure 3.19). The relative weight of each study in the analysis ranged from 14.97 to 70.73%.

Single component vs. multicomponent interventions.

Total SEL. 23 studies used PE, PABAS, or PADS as the single intervention approach to address any SEL outcome. The meta-analysis for these studies indicated a statistically significant and moderate pooled intervention effect on children's total SEL (Hedge's g = 0.47, 95% CI=0.37–0.58; p=.000; I^2 = 69.3%; Figure 3.20). The relative weight of each study in the analysis ranged from 1.75 to 6.92%. Results from the meta-regression found that intervention duration (β = -0.20; 95% CI=-0.42–0.03; p=0.08), delivery agent (β = -0.30; 95% CI=-0.48–0.10; p=0.06), and study design (β = -0.09; 95% CI=-0.30–0.15; p=0.52) were not found to be a statistically significant moderator variables/factors affecting overall study effect sizes (i.e., children's total SEL).

3 studies used multicomponent intervention approaches to address one or more SEL outcome. 1 study included the combination of PADS and PABAS (Bundy et al., 2017), 1 study included the combination of PADS and FCE (Duck et al., 2020), and 1 study included the combination of PE and PABAS (Elbe et al., 2017). The meta-analysis for studies indicated a statistically significant and small pooled intervention effect on children's total SEL (Hedge's g = 0.32, 95% CI=0.13–0.51; p=.001; I^2 = 26.7%; Figure 3.21). The relative weight of each study in the analysis ranged from 12.40 to 20.31%.

DISCUSSION

The primary aim of this systematic review and meta-analysis was to synthesize the evidence of the effectiveness of PA interventions in increasing SEL of children aged 5-12. Thirty-two studies met the inclusion criteria and were included in the qualitative analysis, and twenty-seven studies were included in the quantitative analysis. The CSPAP framework was used to categorize the different intervention approaches and the CASEL

(2020) framework was used to categorize the different intervention outcomes. Considering the aim of this systematic review, the majority of the intervention programs reviewed showed positive effects on students' SEL competence. The results of the aggregate meta-analysis indicate that effect sizes for the development of SEL from pre-post intervention ranged from small to moderate. There is, however, considerable variation in study design, sample size, delivery agent, and study context, and studies were implemented in over 13 countries across diverse settings. The findings of this study suggest that children benefit from receiving CSPAP-aligned PA interventions. That is, exposure to different types of CSPAP single or multiple component intervention approaches results in improvement in the development of children's overall SEL. Further, given the onset of SEL foci in schools, this review can provide key insights into how the CSPAP framework can translate into professional practice.

Results of this review indicated that the majority of studies included PE (n=17) as a component of either a single (89%) or multicomponent (11%) approach and showed beneficial effects on the development of children's SEL. The PE component commonly included the integration of movement-based video games (Andrade et al., 2020; Fu & Burns, 2018; Lwin & Malik, 2014), cognitively challenging movement activities (Cocca et al., 2020; Olive et al., 2019), cooperative games, and used autonomy support strategies for building a positive learning environment as a foundational approach (Moreno-Murcia & Sánchez Latorre, 2016). Some of the included studies found that promotion of SEL in PE lessons affects the enhancement of students' PA levels and FMS, which are important aims of PE (e.g., Annesi et al., 2016a; Chan et al., 2019). A previous study also concluded that a solid foundation can be laid for SEL implementation characterized by

respect, participation, engagement, and cooperation (Wright et al., 2021). Specifically, Owen et al. (2016), based on a meta-analysis of 38 studies, reported a positive association between PA and elements of school engagement, including behavior (e.g., time-on-task), emotions (e.g., lesson enjoyment), and cognition (e.g., self-regulated learning). In addition, our results showed that intervention programs (n=13) included PE components mainly delivered by school-based teams (i.e., PE teachers and classroom teachers). Similarly, Durlak et al. (2011) meta-analyzed 213 school-based SEL programs, including 270,000 students, and demonstrated that teachers and other school staff were able to successfully deliver such programs. Significantly, these findings lend empirical support from a national survey that teachers should play a key role in promoting positive social and emotional growth of students (Bridgeland, Bruce, & Hariharan, 2013). This demonstrates that it is possible to design CSPAP-aligned SEL programs that can become part of routine practice in schools (Jones & Bouffard, 2012).

The majority of included studies (41%) used curriculum-based approaches (e.g., Duck et al., 2020; Fu & Burns, 2018). Further, some studies involved training components for teachers who then implemented CSPAP-aligned PA interventions. The benefit of curriculum-based interventions is that teachers learn implementation, thus eliminating the need for considerable external resources, which is the most commonly identified barrier in school-based efforts (Van den Berg et al., 2017).

From the results, increases in intervention effectiveness may be observed if CSPAP-aligned PA opportunities are designed to specifically integrate the curriculum model for SEL. For instance, some of the included studies show improvement of SEL development by using the TPSR model as a framework in PE lessons and PABAS (Buišić

& Đorđić, 2018; Escartí et al., 2010; Hassandra & Goudas, 2010; PÉRez-OrdÁS et al., 2020; Olive et al., 2020). Specifically, integration of the TPSR model and its key instructional strategies (e.g., awareness talks with teachers, self-reflection, and group meetings for discussion) with intervention programs helped teachers target the development of children's SEL-related outcomes (e.g., personal and social skills). That is, students can learn to discuss and express their emotions (positive or negative), recognize empathy to their peers, and develop conflict resolution skills (Hellison, 2011). This is consistent with previous findings reporting that the TPSR approach within PE lessons was influential in enriching students' personal behaviors (e.g., reduced aggressions and improved caring for others; Cecchini et al., 2007; Pozo et al., 2018). These results and previous literature on the TPSR model, and more broadly on affective learning in PE, suggest that teachers who organized lesson structure using instructional strategies that prioritize the affective domain might have potential to positively change the development of children (Simonton & Shiver, 2021; Webster et al., 2011).

As mentioned above, CSPAP components delivered by school-based teams had a positive effect on the improvement of SEL. In this review, 20 studies (63%) reported PE or classroom teachers' training and support as having 20-30 hour workshop sessions, ongoing professional development meetings, or receiving curriculum resources (e.g., Elbe et al., 2017; Escartí et al., 2010). These studies highlighted that trained teachers who effectively integrated the programs had more positive outcomes with students (e.g., Duck et al., 2020; Grillich et al., 2016; PÉRez-OrdÁS et al., 2020). Unfortunately, SEL professional development is often delivered as a "one-shot" workshop approach, which lacks continuous support for implementation success (Jennings & Frank, 2015).

Therefore, consistent support, goal setting, progress monitoring, and frequent collaborative sessions during which teachers can actively practice new approaches must be included in order for the training to support implementation.

Previous research has shown that several theoretical frameworks have been utilized to provide a conceptual approach to understand and track positive benchmarks of children's SEL development (Evans, Murphy, & Scourfield, 2015; Jones, Barnes, Bailey, & Doolittle, 2017; Lindquist, Reynolds, & Goran, 1999; Wood, 2020). In this review, some studies (30%) cited various theoretical foundations underlying their intervention program. Of these, Social Cognitive Theory (Bandura, 1997) guided five studies (Annesi et al., 2016a, 2016b, 2017; Fu & Burns, 2018; Huang et al., 2018), Achievement Goal Theory (Ames & Archer, 1988) guided one study (Brown & Fry, 2014), Competence Motivation Theory (Harter, 1981) guided one study (Chan et al., 2019), Theory of Planned Behavior (Ajzen, 1991) guided one study (Lwin & Malik, 2014), and Self-Determination Theory (Deci & Ryan, 2012) guided one study (Moreno-Murcia & Sánchez Latorre, 2016). The findings determined that intervention programs have their theoretical roots in approaches that emerged initially as primary strategies to promote the effectiveness of SEL outcomes. For instance, studies grounded in Social Cognitive Theory mainly focused on SA and SM competencies as outcomes within PABAS settings (Annesi et al., 2017; Huang et al., 2018). This could be attributed to developmental implications of the targeting of various social and emotional skills, which include an emphasis on emotional variables in attachment theories (Denham & Burton 2003), social-cognitive variables (in social learning theory), and self-regulation (featuring executive function skills). However, it can be difficult to fully interpret the results of limited studies

in which authors report scant information about the intervention program they investigated. Thus, identifying which theoretical basis (e.g., self-efficacy theory; Bandura, 1971) might be effective for enhancing SEL competencies and coordinating CSPAP-aligned PA programming is important for future research.

Clear evidence is lacking for an ideal intervention dose for improving children's SEL (e.g., intervention duration and frequency). Generally, intervention programs targeted at SEL often take the form of short sessions, implemented during one weekly half-hour or hour-long session (Jones et al., 2010). However, within the included research, limited studies investigating dose were identified, highlighting a clear gap in the literature. It may be worthwhile to consider that the CSPAP-aligned PA intervention programs focusing on the development of SEL had different ranges of duration and time per session, which might have influenced their effectiveness. For instance, based on the results of qualitative and quantitative analysis, some studies reported significant improvements in a single SEL competency for children after a 400-min dose over 10 weeks (Duncan et al., 2018) and an 800-min dose over 8 weeks (Mayorga-Vega et al., 2012), whereas another study showed a relatively small effect on overall SEL competencies after a 720-min dose over 4 weeks (Olive et al., 2020). Other studies showed significant development in children's multiple SEL competencies after a 1,620-min dose over 18 weeks (Fu & Burns, 2018), a 3,240-min dose over 18 weeks (Wong et al., 2017), and a 4,320 min dose over 24-weeks (Annesi et al., 2016a). Overall, although there are unclear timepoints in terms of intervention duration, it is possible that a minimum length of 8 weeks may be needed to bring about positive developmental changes in SEL, which aligns with Teraok et al.'s (2020) findings. Similarly, Wu et al.

(2017) conducted a systematic review of 31 studies and reported that a higher frequency of PA was associated with school-aged children's and youth's better quality of life. Ultimately, understanding patterns of change in SEL that result from different intervention durations could reveal the minimal amount of time needed for positive developmental changes to occur in SEL, and it could help establish recommendations and policies for practitioners implementing CSPAPs. Furthermore, we may consider joint approaches that not only increase quantity aspects (i.e., expanded and extended opportunities for PA) but also quality aspects (i.e., enhanced opportunities for PA; Beets et al., 2016).

When synthesizing the evidence qualitatively, we found considerable heterogeneity in the outcomes, and thus it may be difficult to draw conclusions on effectiveness based on the study characteristics. Furthermore, the wide range of instruments of SEL-related outcomes measured in the studies shows that there is no singular most effective standard. Only a few studies reported the instruments' reliability and validity (e.g., Bundy et al., 2017; Pesce et al., 2012), making it unclear whether the applied measurement method controlled the promoted SEL component in the appropriate manner. In addition, the available data reported in this review are predominantly based on self-report measures, making it necessary to gain further solid evidence in supporting children's SEL. Thus, for future research, a multi-level assessment to be able to measure the SEL-related components will be beneficial and should include observations by peers, teachers, and family via questionnaires as well as objective measurements such as systematic observation by independent observers.

Limited studies performed subgroup analyses (e.g., intervention type, intensity, sample size, and study design). Further, fidelity of implementation and attrition are pertinent factors in considering the effectiveness of SEL programs (Escartí, Llopis-Goig, & Wright, 2018), yet only 38% of included studies explicitly reported using fidelity measures. As a result, we should interpret the findings cautiously due to the lack of precision of the effect size. In parallel, Demitrovitch and Greenberg (2000), in a review of 34 studies, reported that the measure of implementation was usually limited (30%). These limitations restrict the ability to replicate studies that can strengthen the evidence base. In the future, intervention studies should evaluate and report fidelity and intensity of implementation, considering how critical it is for CSPAP-aligned PA programs to operate successfully (Domitrovich et al., 2010; Durlak & DuPre, 2008).

The meta-analysis results indicated that the pooled effect sizes of all PA interventions to develop children's SEL were statistically significant, with moderate effect sizes (Hedge's g=0.44). However, there were a small number of heterogeneous studies included in the meta-analysis. Subsequently, a subgroup comparison between five competencies was performed. Similarly, all included PA intervention programs reported moderate effect sizes on the development of each SEL competency (range from Hedge's g=0.43 to g=0.64). This is in line with Durlak et al.'s (2011) findings of a meta-analysis of school-based, universal SEL programs among K–12 students with similar effect sizes (Cohen's d=0.30). Another systematic review with meta-analysis found that purposefully designed and well-implemented SEL programs are associated with positive outcomes for children, with generally moderate effect sizes (Jones & Bouffard, 2012).

Since mounting evidence supports the idea that PA contributes meaningfully within the area of SEL, practitioners should take a new approach: integrating the PA opportunities and reinforcement of SEL skills into their daily interactions and practices with students (Jones et al., 2012; Olive et al., 2020; Worrell et al., 2020). Research shows that students were more likely to benefit when programs were attentive to places outside classrooms, such as hallways and playgrounds, and were inclusive of parental involvement (Shen et al., 2018; Wong et al., 2016). Thus, research warrants a new perspective and highlights a range of potential approaches and support strategies for more impactful effects of intervention programs.

The current meta-analysis found that both single PE component and single PABAS component interventions positively affect the development of children's overall SEL and competencies, which mainly show moderate effect sizes (range from Hedge's g =0.38 to g=.55). The PE component, within the CSPAP framework, is designed to serve as the primary educational pathway through which students learn the knowledge and skills needed to pursue a lifetime of PA (Centers for Disease Control and Prevention [CDC], 2017; Webster et al., 2020) and is the critical avenue for developing SEL (Ciotto et al., 2018; Wright et al., 2021). In addition, the PABAS component describes a wide variety of PA opportunities (CDC, 2017), which have demonstrated the potential to allow students not only quality minutes of PA outside of PE, but also the improvement of children's SEL (Barcelona, Centeio, & Parker, 2020). Along with this, Greenspan et al. (2019) reviewed that students who engaged in extra moderate-to-vigorous PA opportunities obtained positive outcomes (e.g., increased self-management and cooperation).

Notably, the majority of included studies for the meta-analysis (81%) used PE and/or PABAS components as an intervention approach, whereas no study used SI or FCE components. Similarly, in a review on each individual CSPAP component, Erwin et al. (2013) reported that there is much evidence to support the positive PA outcomes of PE, PABAS, and PADS components. However, the SI component has the potential to greatly support the structure of children's PA and SEL (Prevention CfDCa, 2020). The SI component focuses on the idea that promoting PA should be a collective effort of all members of the school community (CDC, 2019; Webster et al., 2020). For instance, school administrators' support is vital in successfully implementing CSPAP and achieving desired program outcomes (Carson et al., 2014; Orendorff et al., 2020). Further, previous studies also point to the potential of involving school staff to improve SEL (Durlak et al., 2011; Durlak, 2016; Jones et al., 2012). That is, school administrators and staff can take advantage of the power of compositional effects by magnifying and spreading strong SEL competencies and effective SEL practices from one classroom to the school as a whole (Jones & Bouffard, 2012). Moreover, it is important that principals and other administrators attend professional development related to CSPAPs and SEL and communicate that PA and SEL outcomes are priorities for the school. When implementing CSPAPs, some topics that need to be addressed are how to connect the teaching of SEL skills with PA intervention approaches, how to create time and space for program implementation, how to select or develop appropriate PA opportunities, and how to support teachers and staff. Through such an approach, students might participate in PA opportunities together and schools can coordinate with other local organizations to

maximize PA promotion and develop children's SEL (CDC, 2017; Cipriani, Richardson, & Roberts, 2012; Heidorn & Centeio, 2012).

Regardless of the strength of the component, without quality implementation, the potential positive impact of SEL programming is reduced (Durlak et al., 2011). Several studies documented that the success of evidence-based SEL programs depends on high-quality implementation (Domitrovich et al., 2008; Durlak, 2016; Meyers, Durlak, & Wandersman, 2012), which includes following program instructions, following the sequence of activities and lessons, completing most lessons, and using prescribed teaching practices (Durlak & DuPre, 2008; Lendrum & Humphrey, 2012). While some modification for PA programs may be necessary to make practices developmentally appropriate, it is important that teachers and/or researchers understand the purpose and intent of activities as they make adjustments for their classrooms, so that the integrity of the evidence-based program is maintained (Durlak, 2016; Quested et al., 2017).

The current meta-analysis found that single CSPAP component interventions had a larger effect on SEL than multicomponent approaches (Hedge's g=0.47 vs. 0.32). Nevertheless, these findings should be taken with caution, given that only three studies used multicomponent approaches. A possible explanation for this is that it can be more difficult to implement and evaluate the effect of a multicomponent intervention approach on SEL than a single component approach. Regardless, in this review, there is insufficient information on whether more or less CSPAP components enhance children's SEL effectively. More research, therefore, is warranted to examine whether CSPAP-aligned single or multicomponent intervention programs offer greater potential for improving children's SEL as a strategic approach.

Furthermore, one of the potential intervention approaches for maximizing SEL outcomes could be integrating CSPAP components. Specifically, integrating PADS with other components would be the most feasible approach (Pulling Kuhn et al., 2021). Because children spend most of the time in the classroom, it would be logical that interventions combining both PADS (e.g., classroom-based movement integration with classroom teachers; Moon & Webster, 2019) and FCE (e.g., home-based tasks/activities with family) components could be promising ways to develop SEL effectively. Similarly, Murano et al. (2020) found that combining classroom teacher-delivered interventions with involving parents in home settings reported a larger effect size than only classroom teacher interventions for improving children's SEL. This practical notion is also theoretically supported by the ecological systems theory, which views child development as a complex system of relationships affected by multiple levels of the surrounding environment, from immediate settings of family and school to community (Bronfenbrenner, 1986).

In this review, moderation analyses were performed to examine the effectiveness of potential moderators on the effect size of SEL with meta-regression (Hennessy et al., 2019). Specifically, we employed the use of methodological moderator variables, including intervention duration, providers, and study design. In general, our predicted moderators did not show significantly moderate differences in SEL outcomes or improve meta-regression models. Further, a few studies had insignificant heterogeneity, indicating substantial differences in intervention programs, study contexts, and characteristics. In line with the present results, Murano et al. (2020) also reported that the study design, duration, and frequency for intervention were not statistically significant moderators of

effect size. Additionally, there was not enough information available for some moderator variable analyses (e.g., fidelity of implementation, socio-economic status). Thus, it is critical in further studies to collect and report more complete data so that potential moderators can be examined to help further researchers better understand the effects of potential moderator variables on the development of SEL.

Strengths, limitations, and future research

Our research has a number of strengths, including being the first systematic review and meta-analysis of the effectiveness of PA interventions on the development of SEL in high-quality studies in children aged 5-12. Based on the quality assessment, most of the included studies were moderate to strong quality in evidence. The findings can be a useful aid when designing CSPAP-aligned PA interventions, with a range of effective intervention strategy options available in school settings for improving children's SEL. From our results, future research could consider longitudinal SEL outcomes across multiple time points in order to determine the lasting impact of sustainable CSPAP-aligned PA interventions on school-aged children. Another strength of this review was that we used a sensitive search strategy to ensure relevant studies were not missed. In addition, a rigorous review methodology, including independent, duplicate reviews of selected studies, ensured most studies were captured (Hennessy et al., 2019).

This study is not without limitations. First, the SEL outcome measures were not uniform, and there was substantial heterogeneity in their reporting, which makes it difficult to compare findings across included studies. Considering this limitation of, we would recommend that future research related to the development of children's SEL through PA opportunities should be directed to assess standard methods of reporting SEL

outcomes. Second, there was additional substantial heterogeneity among overall effect size estimates, with the PA intervention program accounting for the majority of this variability. There were possible dependency issues in terms of some of the included studies with similar authorship teams, who also implemented similar interventions, which might affect heterogeneity. Thus, their estimate should be interpreted with caution. Third, the review included small sample sizes, and some articles were feasibility studies or pilot trials. For future research, qualitative evidence can help to better understand the 'black-box' of CSPAP-aligned PA program interventions on children's SEL. Finally, for practical reasons, our review was limited to published research in the English language and excluded unpublished literature and dissertations, potentially producing bias into the study results. We may have therefore missed unpublished work and/or literature published in other languages.

Implication for research and practice

We recommend that those wishing to implement CSPAP-aligned PA interventions targeting development of children's SEL use high-quality, rigorously evaluated, and age-specific evidence in selecting or creating a developmentally appropriate program that will benefit students. McClelland et al. (2017, p.39) suggest that various factors influence intervention effectiveness, and "a one-size-fits-all approach to intervention may not help all children." The developmental period of 5-12 years of age are too critical for educators not to select curricula that will confer the greatest benefits (Goodman et al., 2015). One way to do this would be to link CSPAP approaches with the SAFE principles recommended as best practices for SEL intervention programs (Durlak et al., 2011): Sequenced: The program has a coordinated progression of activities or

practices to build competencies; Active: Participatory elements (e.g., role plays) include students in active forms of learning; Focused: There is a dedicated time or specific program element that is focused on developing SEL competencies; and Explicit: The program explicitly targets SEL competencies that it is trying to develop within the intervention. SEL programs using all four of the SAFE principles had positive effects on more outcomes than programs using less than all four of these principles (Durlak et al., 2011). Through such an approach, understanding how CSPAP affects children's outcomes could enhance advocacy efforts among administration and school staff who prioritize SEL development, leading to positive changes in school culture and policies for PA.

CONCLUSION

Building on the benefits that many studies have documented, the results of this systematic review and meta-analysis provide evidence that PA intervention programs that align with one or more components of the CSPAP framework can foster SEL competencies. This study adds to a growing body of SEL research within the CSPAP framework by examining the potential pathways between different types of PA intervention approaches which aid children's SEL development. Achieving optimal SEL in children is an essential factor for their healthy development and well-being in future stages of life (Greenberg et al., 2017). More studies are needed to advance this research agenda. Large sample sizes will ensure a study is appropriately powered and generalizable and rigorous randomized designs will ensure more confidence in making causal claims. Practitioners, administrators, scholars, and policymakers should consider

the unique needs of the school-aged children carefully and use CSPAP-aligned PA intervention programs to promote children's SEL successfully.

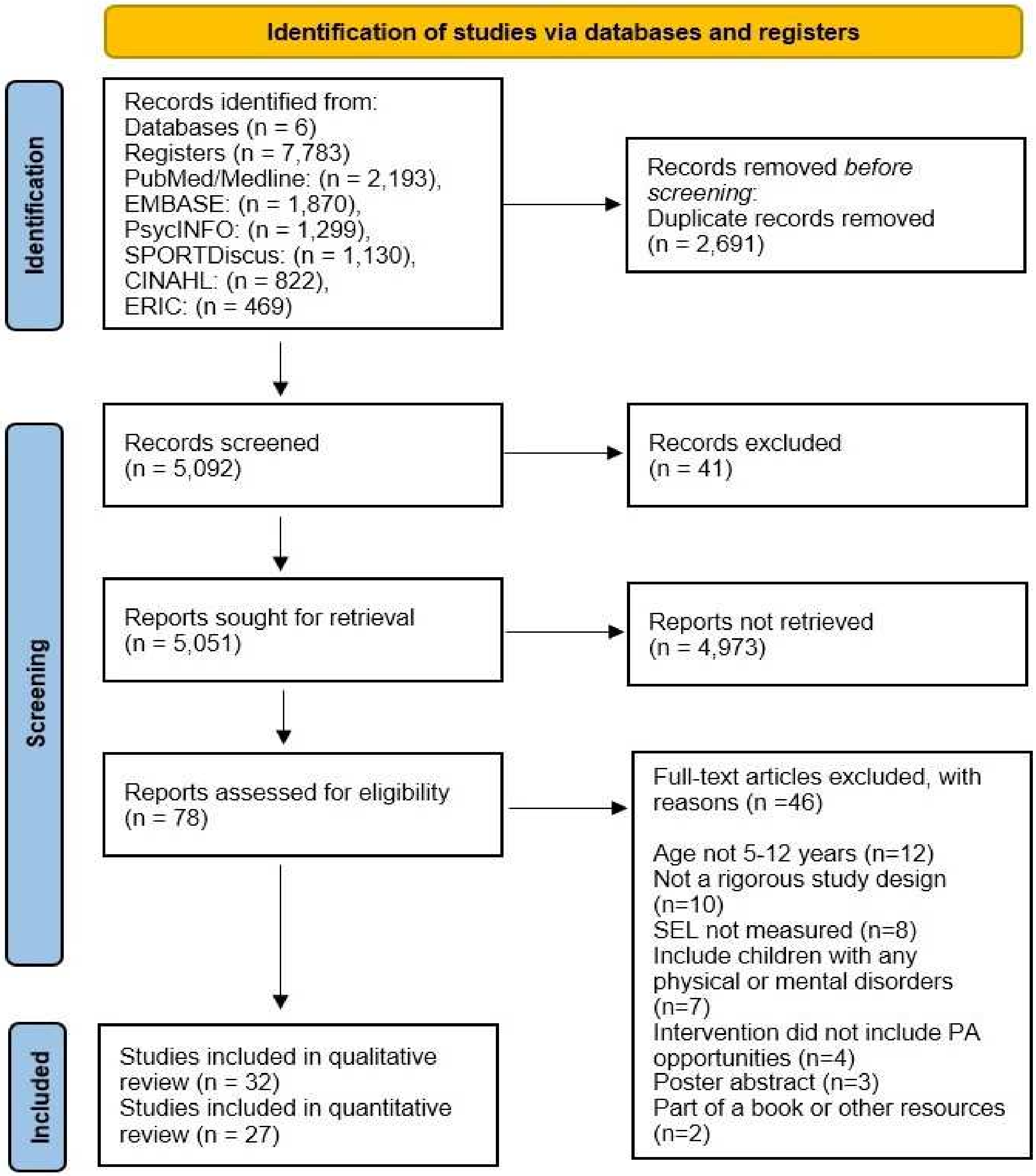

Figure 3.1 PRISMA flow chart of the search process of screened, included, and excluded articles.

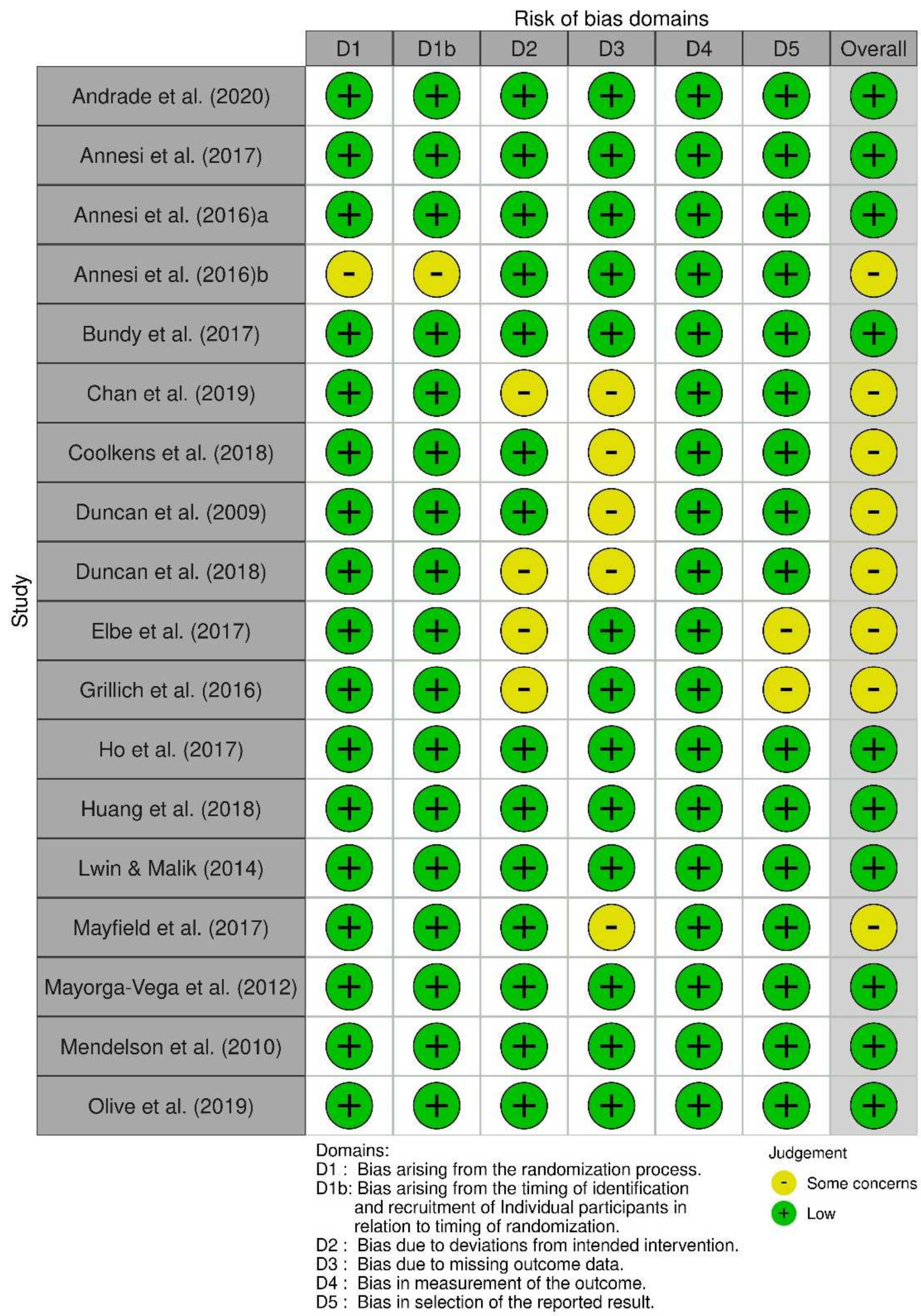

Figure 3.2 Quality assessment for C-RCT or RCT studies based on Revised Cochrane Risk of Bias tool for randomized trials (RoB 2.0).

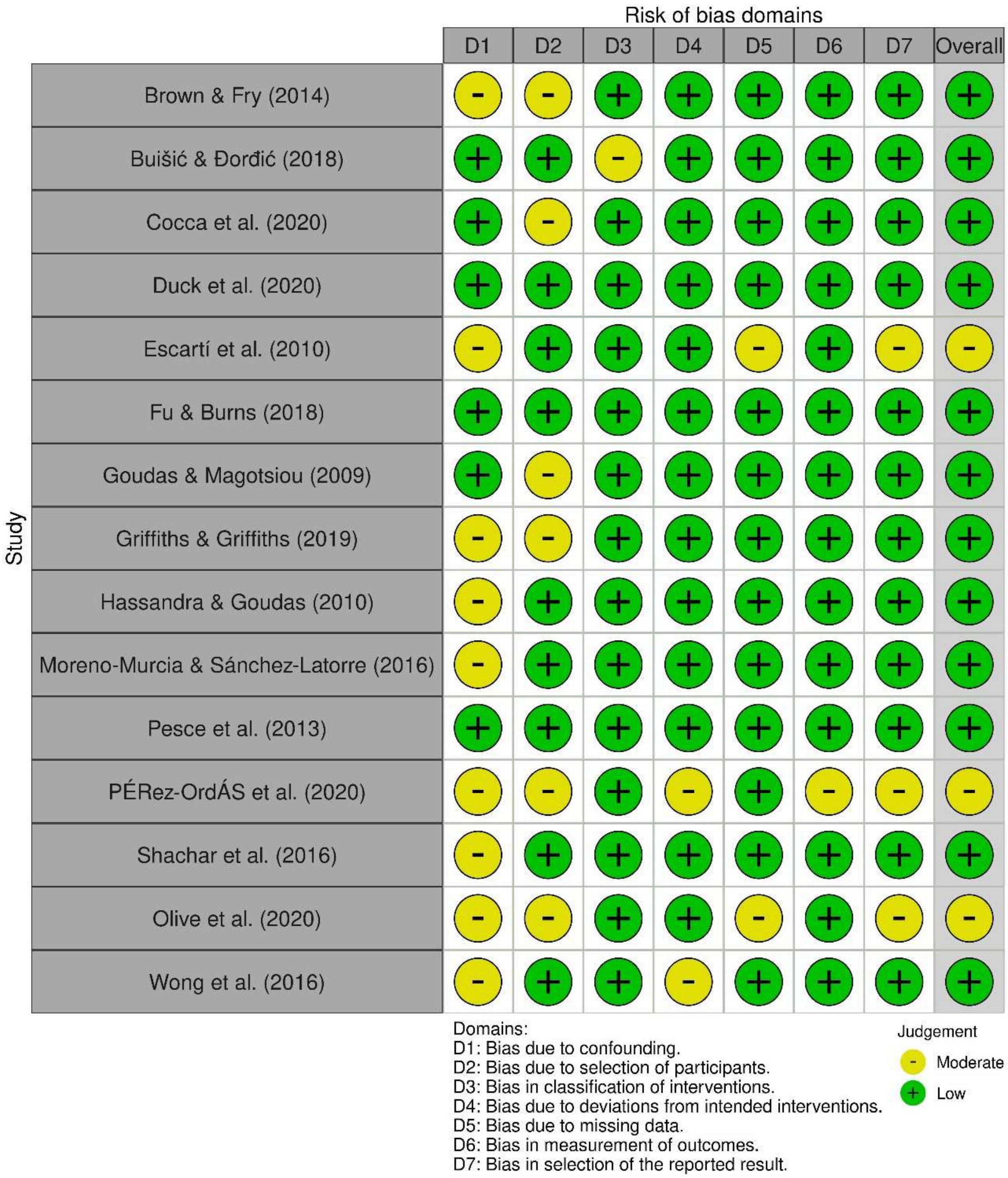

Figure 3.3 Quality assessment for N-RCT studies based on Risk of Bias in Non-randomized Studies of Interventions tool (ROBINS-I).

Study name	Hedges's g	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
Andrade et al. 2020	0.584	0.172	0.030	0.247	0.920	3.397	0.001
Annesi et al. 2016a	0.218	0.170	0.029	-0.115	0.550	1.283	0.199
Annesi et al. 2016b	0.992	0.179	0.032	0.642	1.343	5.548	0.000
Annesi et al. 2017	0.445	0.174	0.030	0.105	0.786	2.562	0.010
Brown & Fry 2014	0.545	0.153	0.023	0.246	0.844	3.571	0.000
Buisic et al. 2018	0.901	0.209	0.044	0.491	1.312	4.304	0.000
Bundy et al. 2017	0.126	0.137	0.019	-0.143	0.396	0.921	0.357
Chan et al. 2019	0.337	0.123	0.015	0.096	0.578	2.740	0.006
Cocca et al. 2020	0.106	0.133	0.018	-0.154	0.366	0.798	0.425
Duck et al. 2020	0.330	0.333	0.111	-0.322	0.982	0.991	0.322
Duncan et al. 2009	0.621	0.246	0.060	0.140	1.103	2.529	0.011
Duncan et al. 2018	0.700	0.211	0.045	0.286	1.113	3.316	0.001
Elbe et al. 2017	0.156	0.144	0.021	-0.126	0.437	1.082	0.279
Escarti al. 2010	0.338	0.305	0.093	-0.260	0.936	1.109	0.267
Fu et al. 2018	0.921	0.258	0.067	0.415	1.426	3.567	0.000
Goudas et al. 2009	0.241	0.187	0.035	-0.125	0.607	1.291	0.197
Griffiths & Griffiths 2019	0.117	0.101	0.010	-0.080	0.314	1.164	0.244
Hassandra et al. 2010	0.439	0.310	0.096	-0.169	1.047	1.416	0.157
Ho et al. 2017	0.293	0.078	0.006	0.141	0.446	3.765	0.000
Lwin & Malik 2014	0.429	0.101	0.010	0.230	0.628	4.232	0.000
Mayorga-Vega et al. 2012	1.119	0.256	0.066	0.616	1.622	4.364	0.000
Moreno-Murcia et al. 2016	0.501	0.173	0.030	0.161	0.841	2.889	0.004
Olive et al. 2019	0.334	0.043	0.002	0.250	0.419	7.800	0.000
Olive et al. 2020	0.095	0.367	0.134	-0.624	0.814	0.259	0.796
Pesce et al. 2013	0.323	0.179	0.032	-0.028	0.673	1.802	0.072
Wong et al. 2016	0.756	0.083	0.007	0.593	0.918	9.112	0.000
	0.437	0.051	0.003	0.337	0.537	8.572	0.000

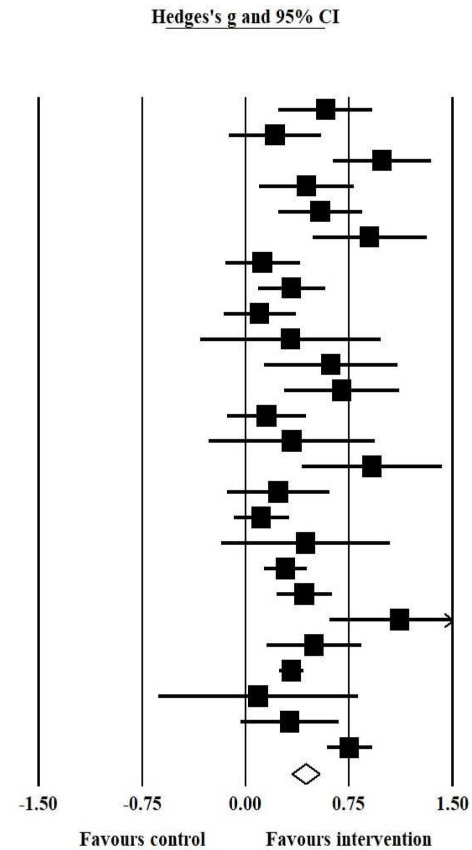

Figure 3.4. Forest plot for pooled effect sizes (Hedge's g) of all CSPAP-aligned PA interventions on children's total SEL from random effects meta-analysis of eligible studies.

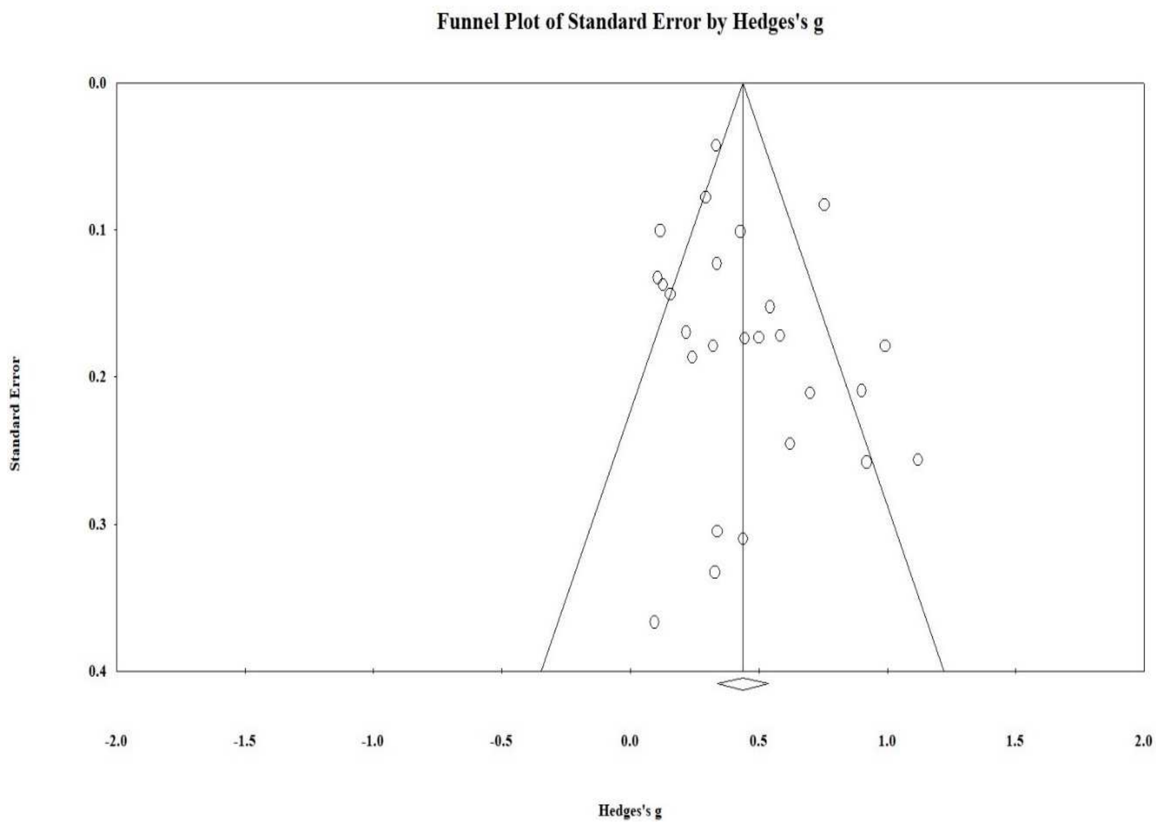

Figure 3.5 Funnel plot for assessment of publication bias using pooled analysis of effect sizes of SEL outcomes

Study name	Hedges's g	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
Andrade et al. 2020	0.342	0.169	0.029	0.010	0.674	2.016	0.044
Annesi et al. 2016a	0.408	0.171	0.029	0.073	0.742	2.386	0.017
Annesi et al. 2016b	0.992	0.179	0.032	0.642	1.343	5.548	0.000
Annesi et al. 2017	0.445	0.174	0.030	0.105	0.786	2.562	0.010
Bundy et al. 2017	0.174	0.136	0.019	-0.094	0.441	1.273	0.203
Chan et al. 2019	0.369	0.123	0.015	0.127	0.610	2.992	0.003
Cocca et al. 2020	0.271	0.133	0.018	0.010	0.532	2.034	0.042
Duncan et al. 2009	0.621	0.246	0.060	0.140	1.103	2.529	0.011
Duncan et al. 2018	0.700	0.211	0.045	0.286	1.113	3.316	0.001
Elbe et al. 2017	0.054	0.144	0.021	-0.228	0.335	0.375	0.708
Escarti et al. 2010	0.650	0.311	0.097	0.041	1.260	2.091	0.036
Fu et al. 2018	0.724	0.253	0.064	0.227	1.220	2.858	0.004
Griffiths & Griffiths 2019	0.359	0.101	0.010	0.161	0.557	3.555	0.000
Hassandra et al. 2010	0.628	0.314	0.099	0.013	1.244	2.000	0.046
Ho et al. 2017	0.409	0.078	0.006	0.255	0.562	5.219	0.000
Lwin & Malik 2014	0.563	0.102	0.010	0.363	0.763	5.510	0.000
Mayorga-Vega et al. 2012	0.931	0.251	0.063	0.439	1.423	3.710	0.000
Moreno Murcia 2016	0.473	0.173	0.030	0.134	0.812	2.733	0.006
Olive et al. 2019	0.244	0.043	0.002	0.160	0.328	5.706	0.000
Pesce et al. 2013	0.465	0.180	0.032	0.112	0.818	2.581	0.010
Wong et al. 2016	0.413	0.081	0.007	0.254	0.572	5.093	0.000
	0.429	0.045	0.002	0.341	0.518	9.547	0.000

Statistics for each study

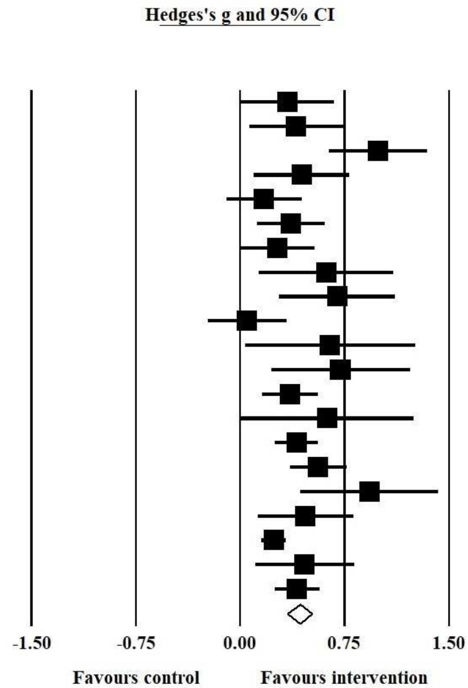

Figure 3.6 Forest plot for pooled effect sizes (Hedge's g) of all CSPAP-aligned PA interventions on children's SA from random effects meta-analysis of eligible studies.

Study name	Hedges's g	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
Andrade et al. 2020	0.608	0.172	0.030	0.271	0.945	3.534	0.000
Annesi et al. 2016a	0.101	0.169	0.029	-0.230	0.433	0.598	0.550
Annesi et al. 2017	0.532	0.175	0.030	0.190	0.874	3.045	0.002
Buisic et al. 2018	0.467	0.202	0.041	0.071	0.864	2.312	0.021
Cocca et al. 2020	0.036	0.133	0.018	-0.224	0.296	0.273	0.785
Escarti et al. 2010	0.332	0.305	0.093	-0.266	0.930	1.088	0.276
Fu et al. 2018	0.821	0.255	0.065	0.320	1.322	3.214	0.001
Goudas et al. 2009	0.606	0.190	0.036	0.233	0.979	3.184	0.001
Griffiths et al. 2019	0.145	0.101	0.010	-0.052	0.342	1.444	0.149
Ho et al. 2017	0.257	0.078	0.006	0.104	0.409	3.300	0.001
Moreno-Murcia et al. 2016	0.829	0.178	0.032	0.481	1.178	4.667	0.000
Olive et al. 2019	0.132	0.043	0.002	0.048	0.215	3.095	0.002
	0.357	0.072	0.005	0.216	0.497	4.965	0.000

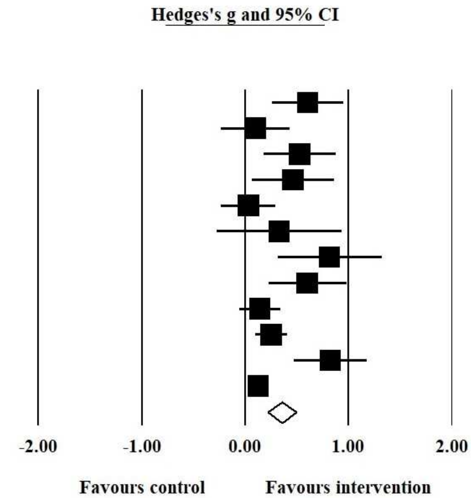

Figure 3.7 Forest plot for pooled effect sizes (Hedge's g) of all CSPAP-aligned PA interventions on children's SM from random effects meta-analysis of eligible studies.

Study name	Hedges's g	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
Buisic et al. 2018	0.860	0.209	0.043	0.451	1.269	4.124	0.000
Chan et al. 2019	0.336	0.123	0.015	0.095	0.577	2.735	0.006
Escarti et al. 2010	0.404	0.306	0.094	-0.196	1.004	1.320	0.187
Goudas et al. 2009	0.245	0.187	0.035	-0.121	0.611	1.311	0.190
Griffiths et al. 2019	0.174	0.101	0.010	-0.023	0.372	1.731	0.084
Lwin et al. 2014	0.429	0.101	0.010	0.231	0.628	4.235	0.000
Olive et al. 2019	0.436	0.043	0.002	0.351	0.520	10.109	0.000
Pesce et al. 2013	0.356	0.179	0.032	0.004	0.707	1.984	0.047
Wong et al. 2016	0.772	0.083	0.007	0.609	0.935	9.297	0.000
	0.444	0.073	0.005	0.301	0.587	6.067	0.000

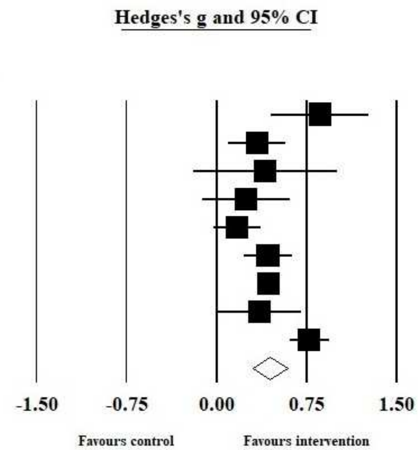

Figure 3.8 Forest plot for pooled effect sizes (Hedge's g) of all CSPAP-aligned PA interventions on children's SoA from random effects meta-analysis of eligible studies.

Study name	Statistics for each study						
	Hedges's g	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
Brown & Fry 2014	0.648	0.157	0.025	0.341	0.955	4.135	0.000
Buisic et al. 2018	1.058	0.213	0.045	0.641	1.476	4.968	0.000
Bundy et al. 2017	0.220	0.138	0.019	-0.050	0.489	1.597	0.110
Duck et al. 2020	0.385	0.334	0.111	-0.269	1.039	1.153	0.249
Elbe et al. 2017	0.156	0.144	0.021	-0.126	0.438	1.083	0.279
Fu et al. 2018	1.298	0.270	0.073	0.768	1.827	4.801	0.000
Goudas et al. 2009	0.245	0.187	0.035	-0.121	0.611	1.313	0.189
Griffiths et al. 2019	0.168	0.101	0.010	-0.029	0.365	1.668	0.095
Ho et al. 2017	0.306	0.078	0.006	0.153	0.459	3.924	0.000
Moreno-Murcia et al. 2016	0.578	0.174	0.030	0.237	0.920	3.320	0.001
Olive et al. 2019	0.429	0.043	0.002	0.344	0.513	9.952	0.000
	0.437	0.075	0.006	0.291	0.583	5.870	0.000

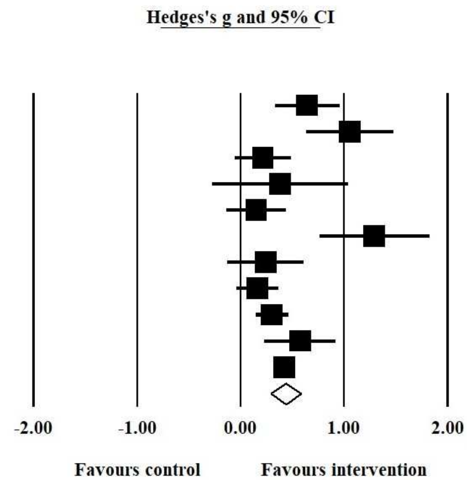

Figure 3.9 Forest plot for pooled effect sizes (Hedge's g) of all CSPAP-aligned PA interventions on children's RS from random effects meta-analysis of eligible studies.

Study name	Hedges's g	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
Brown & Fry 2014	0.683	0.158	0.025	0.373	0.993	4.318	0.000
Buisic et al. 2018	0.780	0.207	0.043	0.375	1.186	3.771	0.000
Hassandra et al. 2010	0.492	0.311	0.097	-0.118	1.102	1.581	0.114
Pesce et al. 2013	0.403	0.180	0.032	0.051	0.755	2.245	0.025
Wong et al. 2016	0.664	0.082	0.007	0.503	0.826	8.071	0.000
	0.639	0.063	0.004	0.516	0.762	10.147	0.000

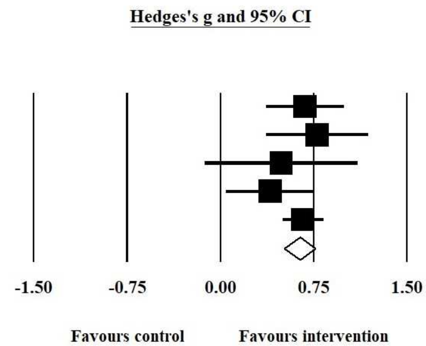

Figure 3.10 Forest plot for pooled effect sizes (Hedge's g) of all CSPAP-aligned PA interventions on children's RDM from random effects meta-analysis of eligible studies.

Study name	Statistics for each study						
	Hedges's g	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
Andrade et al. 2020	0.584	0.172	0.030	0.247	0.920	3.397	0.001
Buisic et al. 2018	0.901	0.209	0.044	0.491	1.312	4.304	0.000
Chan et al. 2019	0.337	0.123	0.015	0.096	0.578	2.740	0.006
Cocca et al. 2020	0.106	0.133	0.018	-0.154	0.366	0.798	0.425
Duncan et al. 2018	0.700	0.211	0.045	0.286	1.113	3.316	0.001
Escarti al. 2010	0.338	0.305	0.093	-0.260	0.936	1.109	0.267
Goudas et al. 2009	0.241	0.187	0.035	-0.125	0.607	1.291	0.197
Griffiths & Griffiths 2019	0.117	0.101	0.010	-0.080	0.314	1.164	0.244
Hassandra et al. 2010	0.439	0.310	0.096	-0.169	1.047	1.416	0.157
Lwin & Malik 2014	0.429	0.101	0.010	0.230	0.628	4.232	0.000
Mayorga-Vega et al. 2012	1.119	0.256	0.066	0.616	1.622	4.364	0.000
Moreno-Murcia et al. 2016	0.501	0.173	0.030	0.161	0.841	2.889	0.004
Olive et al. 2019	0.334	0.043	0.002	0.250	0.419	7.800	0.000
Pesce et al. 2013	0.323	0.179	0.032	-0.028	0.673	1.802	0.072
	0.407	0.060	0.004	0.290	0.525	6.786	0.000

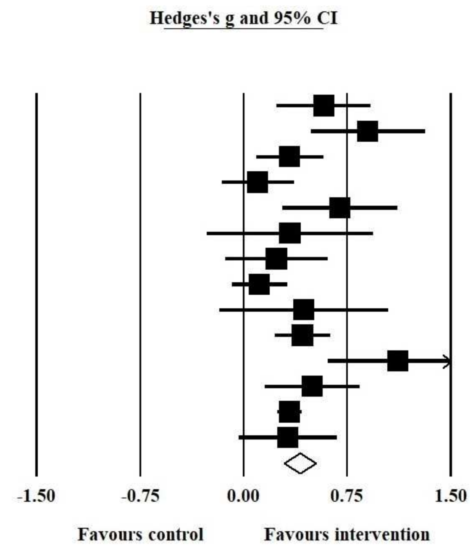

Figure 3.11 Forest plot for pooled effect sizes (Hedge’s g) of PE single component interventions on children’s total SEL from random effects meta-analysis of eligible studies.

Study name	Hedges's g	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
Andrade et al. 2020	0.342	0.169	0.029	0.010	0.674	2.016	0.044
Chan et al. 2019	0.369	0.123	0.015	0.127	0.610	2.992	0.003
Cocca et al. 2020	0.271	0.133	0.018	0.010	0.532	2.034	0.042
Duncan et al. 2018	0.700	0.211	0.045	0.286	1.113	3.316	0.001
Escarti et al. 2010	0.650	0.311	0.097	0.041	1.260	2.091	0.036
Griffiths & Griffiths 2019	0.359	0.101	0.010	0.161	0.557	3.555	0.000
Hassandra et al. 2010	0.628	0.314	0.099	0.013	1.244	2.000	0.046
Lwin & Malik 2014	0.563	0.102	0.010	0.363	0.763	5.510	0.000
Mayorga-Vega et al. 2012	0.931	0.251	0.063	0.439	1.423	3.710	0.000
Moreno Murcia 2016	0.473	0.173	0.030	0.134	0.812	2.733	0.006
Olive et al. 2019	0.244	0.043	0.002	0.160	0.328	5.706	0.000
Pesce et al. 2013	0.465	0.180	0.032	0.112	0.818	2.581	0.010
	0.426	0.056	0.003	0.316	0.537	7.562	0.000

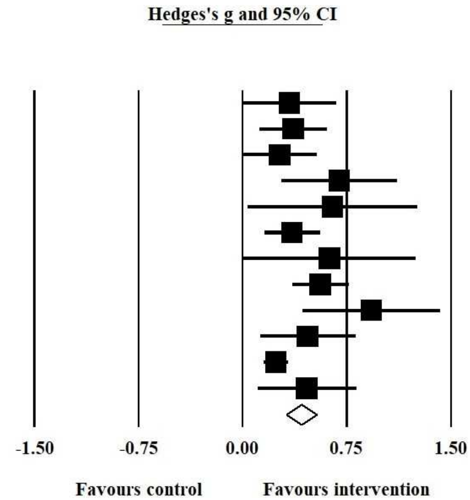

Figure 3.12 Forest plot for pooled effect sizes (Hedge's g) of PE single component interventions on children's SA from random effects meta-analysis of eligible studies.

Study name	Hedges's g	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
Andrade et al. 2020	0.608	0.172	0.030	0.271	0.945	3.534	0.000
Buisic et al. 2018	0.476	0.202	0.041	0.079	0.872	2.351	0.019
Cocca et al. 2020	0.338	0.133	0.018	0.077	0.600	2.535	0.011
Escarti et al. 2010	0.332	0.305	0.093	-0.266	0.930	1.088	0.276
Goudas et al. 2009	0.606	0.190	0.036	0.233	0.979	3.184	0.001
Griffiths et al. 2019	0.145	0.101	0.010	-0.052	0.342	1.444	0.149
Moreno-Murcia et al. 2016	0.829	0.178	0.032	0.481	1.178	4.667	0.000
Olive et al. 2019	0.217	0.043	0.002	0.134	0.301	5.086	0.000
	0.406	0.084	0.007	0.242	0.570	4.856	0.000

Statistics for each study

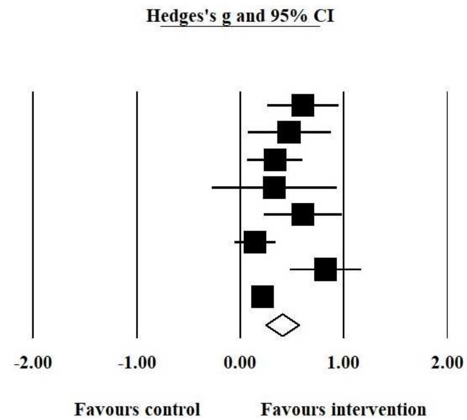

Figure 3.13 Forest plot for pooled effect sizes (Hedge’s g) of PE single component interventions on children’s SM from random effects meta-analysis of eligible studies.

Study name	Statistics for each study						
	Hedges's g	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
Buisic et al. 2018	0.860	0.209	0.043	0.451	1.269	4.124	0.000
Chan et al. 2019	0.336	0.123	0.015	0.095	0.577	2.735	0.006
Escarti et al. 2010	0.404	0.306	0.094	-0.196	1.004	1.320	0.187
Goudas et al. 2009	0.245	0.187	0.035	-0.121	0.611	1.311	0.190
Griffiths et al. 2019	0.174	0.101	0.010	-0.023	0.372	1.731	0.084
Lwin et al. 2014	0.429	0.101	0.010	0.231	0.628	4.235	0.000
Olive et al. 2019	0.436	0.043	0.002	0.351	0.520	10.109	0.000
Pesce et al. 2013	0.356	0.179	0.032	0.004	0.707	1.984	0.047
	0.383	0.056	0.003	0.273	0.492	6.843	0.000

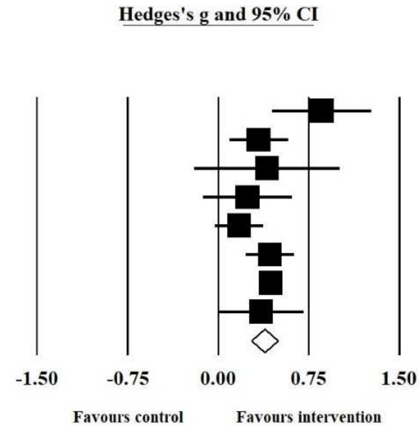

Figure 3.14 Forest plot for pooled effect sizes (Hedge’s g) of PE single component interventions on children’s SoA from random effects meta-analysis of eligible studies.

Study name	Hedges's g	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
Buisic et al. 2018	1.058	0.213	0.045	0.641	1.476	4.968	0.000
Goudas et al. 2009	0.245	0.187	0.035	-0.121	0.611	1.313	0.189
Griffiths et al. 2019	0.168	0.101	0.010	-0.029	0.365	1.668	0.095
Moreno-Murcia et al. 2016	0.578	0.174	0.030	0.237	0.920	3.320	0.001
Olive et al. 2019	0.429	0.043	0.002	0.344	0.513	9.952	0.000
	0.453	0.112	0.013	0.232	0.673	4.025	0.000

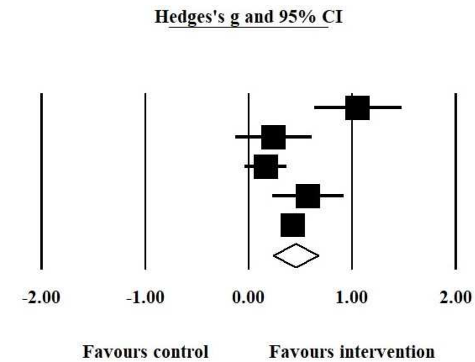

Figure 3.15 Forest plot for pooled effect sizes (Hedge’s g) of PE single component interventions on children’s RS from random effects meta-analysis of eligible studies.

Study name	Hedges's g	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
Buisic et al. 2018	0.780	0.207	0.043	0.375	1.186	3.771	0.000
Hassandra et al. 2010	0.492	0.311	0.097	-0.118	1.102	1.581	0.114
Pesce et al. 2013	0.403	0.180	0.032	0.051	0.755	2.245	0.025
	0.554	0.124	0.015	0.310	0.797	4.452	0.000

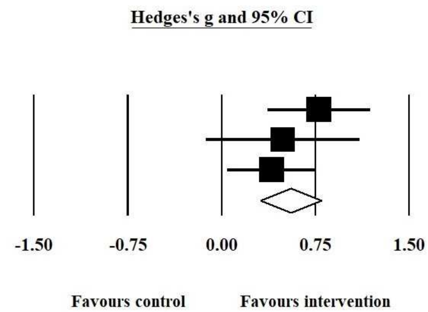

Figure 3.16 Forest plot for pooled effect sizes (Hedge's g) of PE single component interventions on children's RDM from random effects meta-analysis of eligible studies.

Study name	Hedges's g	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
	Statistics for each study						
Annesi et al. (2016)a	0.218	0.170	0.029	-0.115	0.550	1.283	0.199
Annesi et al. (2016)b	0.992	0.179	0.032	0.642	1.343	5.548	0.000
Annesi et al. (2017)	0.445	0.174	0.030	0.105	0.786	2.562	0.010
Brown & Fry (2014)	0.545	0.153	0.023	0.246	0.844	3.571	0.000
Duncan et al. (2009)	0.621	0.246	0.060	0.140	1.103	2.529	0.011
Ho et al. (2017)	0.293	0.078	0.006	0.141	0.446	3.765	0.000
Olive et al. (2020)	0.095	0.367	0.134	-0.624	0.814	0.259	0.796
Wong et al. (2016)	0.756	0.083	0.007	0.593	0.918	9.112	0.000
	0.520	0.105	0.011	0.313	0.726	4.928	0.000

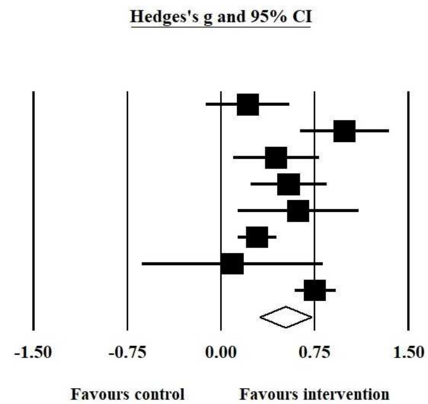

Figure 3.17 Forest plot for pooled effect sizes (Hedge’s g) of PABAS single component interventions on children’s total SEL from random effects meta-analysis of eligible studies.

Study name	Hedges's g	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
	Statistics for each study						
Annesi et al. 2016a	0.408	0.171	0.029	0.073	0.742	2.386	0.017
Annesi et al. 2016b	0.992	0.179	0.032	0.642	1.343	5.548	0.000
Annesi et al. 2017	0.445	0.174	0.030	0.105	0.786	2.562	0.010
Brown et al. 2014	0.606	0.155	0.024	0.303	0.910	3.913	0.000
Duncan et al. 2009	0.870	0.251	0.063	0.378	1.362	3.463	0.001
Ho et al. 2017	0.409	0.078	0.006	0.255	0.562	5.219	0.000
Wong et al. 2016	0.507	0.081	0.007	0.347	0.667	6.225	0.000
	0.556	0.073	0.005	0.413	0.699	7.622	0.000

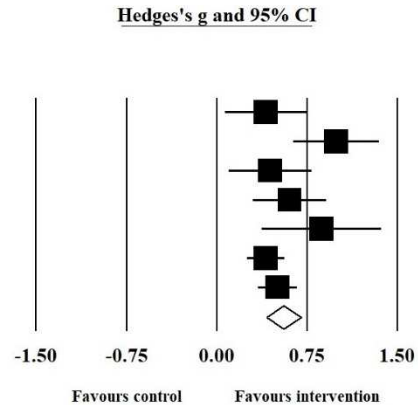

Figure 3.18 Forest plot for pooled effect sizes (Hedge's g) of PABAS single component interventions on children's SA from random effects meta-analysis of eligible studies.

Study name	Hedges's g	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
Annesi et al. 2016a	0.103	0.169	0.029	-0.229	0.435	0.608	0.543
Annesi et al. 2017	0.366	0.173	0.030	0.027	0.705	2.114	0.035
Ho et al. 2017	0.257	0.078	0.006	0.104	0.409	3.300	0.001
	0.249	0.065	0.004	0.121	0.378	3.810	0.000

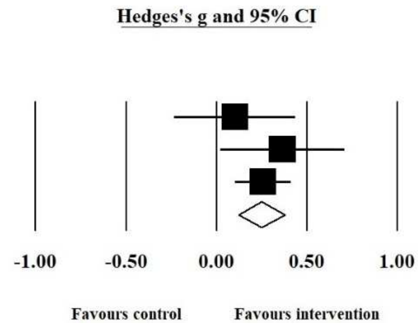

Figure 3.19 Forest plot for pooled effect sizes (Hedge's g) of PABAS single component interventions on children's SM from random effects meta-analysis of eligible studies.

Study name	Hedges's g	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
	Statistics for each study						
Andrade et al. 2020	0.584	0.172	0.030	0.247	0.920	3.397	0.001
Annesi et al. 2016a	0.218	0.170	0.029	-0.115	0.550	1.283	0.199
Annesi et al. 2016b	0.992	0.179	0.032	0.642	1.343	5.548	0.000
Annesi et al. 2017	0.445	0.174	0.030	0.105	0.786	2.562	0.010
Brown & Fry 2014	0.545	0.153	0.023	0.246	0.844	3.571	0.000
Buisic et al. 2018	0.901	0.209	0.044	0.491	1.312	4.304	0.000
Chan et al. 2019	0.337	0.123	0.015	0.096	0.578	2.740	0.006
Cocca et al. 2020	0.106	0.133	0.018	-0.154	0.366	0.798	0.425
Duncan et al. 2009	0.621	0.246	0.060	0.140	1.103	2.529	0.011
Duncan et al. 2018	0.700	0.211	0.045	0.286	1.113	3.316	0.001
Escarti al. 2010	0.338	0.305	0.093	-0.260	0.936	1.109	0.267
Fu & Burns 2018	0.921	0.258	0.067	0.415	1.426	3.567	0.000
Goudas et al. 2009	0.241	0.187	0.035	-0.125	0.607	1.291	0.197
Griffiths & Griffiths 2019	0.117	0.101	0.010	-0.080	0.314	1.164	0.244
Hassandra et al. 2010	0.439	0.310	0.096	-0.169	1.047	1.416	0.157
Ho et al. 2017	0.293	0.078	0.006	0.141	0.446	3.765	0.000
Lwin & Malik 2014	0.429	0.101	0.010	0.230	0.628	4.232	0.000
Mayorga-Vega et al. 2012	1.119	0.256	0.066	0.616	1.622	4.364	0.000
Moreno-Murcia et al. 2016	0.501	0.173	0.030	0.161	0.841	2.889	0.004
Olive et al. 2019	0.095	0.367	0.134	-0.624	0.814	0.259	0.796
Olive et al. 2020	0.334	0.043	0.002	0.250	0.419	7.800	0.000
Pesce et al. 2013	0.323	0.179	0.032	-0.028	0.673	1.802	0.072
Wong et al. 2016	0.756	0.083	0.007	0.593	0.918	9.112	0.000
	0.470	0.054	0.003	0.363	0.576	8.655	0.000

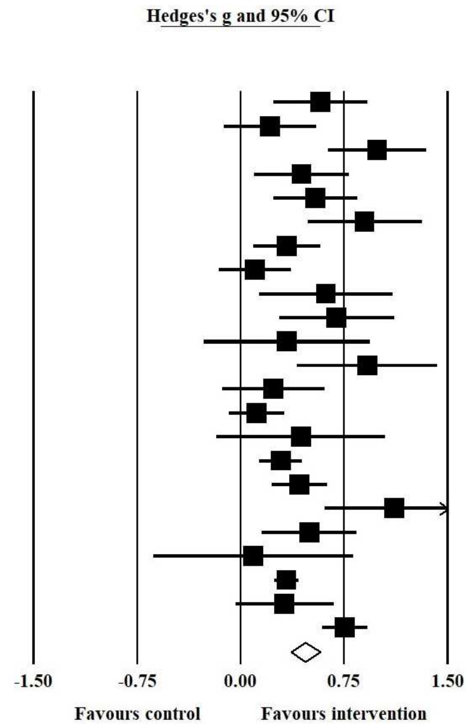

Figure 3.20 Forest plot for pooled effect sizes (Hedge's g) of single component interventions on children's total SEL from random effects meta-analysis of eligible studies.

Study name	Hedges's g	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
Bundy et al. 2017	0.523	0.140	0.019	0.250	0.797	3.751	0.000
Duck et al. 2020	0.330	0.333	0.111	-0.322	0.982	0.991	0.322
Elbe et al. 2017	0.102	0.144	0.021	-0.179	0.384	0.712	0.477
	0.320	0.096	0.009	0.132	0.508	3.337	0.001

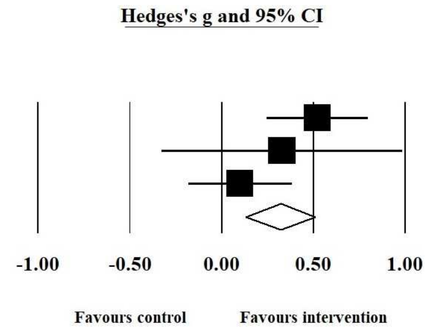

Figure 3.21 Forest plot for pooled effect sizes (Hedge's g) of multiple component interventions on children's total SEL from random effects meta-analysis of eligible studies.

CHAPTER 4

STUDY 3: DEVELOPMENT OF THE SYSTEM FOR OBSERVING VIRTUAL REAL TIME LESSONS IN PHYSICAL EDUCATION (SOVRTL-PE)

INTRODUCTION

In 2007, the National Association for Sports and Physical Education (NASPE) reported that physical education (PE) teachers were divided on their perspective of the viability of online physical education (OLPE; NASPE, 2007). Some have continued to question the validity of synchronous OLPE as a viable educational alternative for face-to-face PE due to a perceived negative impact on students' learning (Mosier, 2012; Ransdell, et al., 2008; Rhea, 2011). Correspondingly, Buschner (2006), notes that without relevant, evidence-based teaching practices/instructions, "OLPE is still only a box that involves sophisticated sound, images, lights, and words that purport to help students learn about and be physically active—but it is not PE in its present mode" (p. 5). However, online education has experienced tremendous growth over the past decade, and, currently, 5.8 million high school students are predicted to be enrolled in online courses (Allen & Seaman, 2016; Gemin & Pape, 2016). Alongside this expansion, OLPE has become more prevalent in the United States during the past decade, with 32 states now allowing students to complete PE credits online (Digital Learning Collaborative [DLC], 2019).

Further, due to the COVID-19 pandemic, both face-to-face school attendance and online classes are occurring in tandem around the world (DLC, 2019, 2020).

Limited research has been performed to examine whether synchronous OLPE lessons are being held and if they are conveying the quality of PE lessons. Through OLPE, the PE national content standards for K-12 learners should be addressed, but with different instructional approaches (Society of Health and Physical Educators [SHAPE] America, 2015, 2018). Prior studies show the efficiency and potential of OLPE classes (Daum, 2020; Daum & Woods, 2015; Mohnsen, 2012). However, other studies indicate the limited nature of OLPE programs, arguing they are solely fitness-based and not comprehensive PE programs inclusive of knowledge-, social- and cooperative skill development (Daum & Buschner, 2012; Daum & Woods, 2015). The nature of OLPE lessons requires special preparation and operation to communicate and practice quality PE (Flores, 2020; SHAPE America, 2018). However, little is known about the nature of synchronous OLPE teaching practices. Research to date is limited to qualitative approaches and is devoid of objective accounts derived from systematic observation (Kooiman et al. 2017; Williams et al., 2020). Additionally, with out-of-date guidelines and a lack of established best practices for OLPE, developed curricula have lacked consistency. Without relevant research into best practices, teachers are finding it difficult to expand the curriculum past the cognitive and fitness aspects of PE (Howley, 2021; Mercier et al., 2021). Thus, there is a need to identify the existing practices of OLPE and establish the best directions for future OLPE programming.

Synchronous OLPE, which offers the closest virtual parallel to in-person classroom instruction, needs further investigation to inform best practice guidelines

regarding teaching strategies for creating quality and meaningful learning experiences for students. Further, there is no way to systematically measure and assess “good” or “bad” synchronous OLPE instruction; conversely, there are numerous established forms of measurement for in-person PE classes (e.g., Rink & Werner, 1989; Roberts & Fairclough, 2012), which presents an opportunity to aid the development of such systems for synchronous OLPE instruction. In addition, the comprehensive school physical activity program (CSPAP) framework encompasses multiple components through which PE can be delivered and supported, including a component that focuses on family and community engagement (Centers for Disease Control and Prevention [CDC], 2019). However, little research has focused on how to include and optimize family and community engagement resources in order to support students’ learning. Given that support from families and communities may be essential to effectively implement PE that is delivered virtually (Webster et al., 2021), there is a need to investigate how to use family and community resources to enhance students’ OLPE learning outcomes.

The purpose of this study was to provide an overview of the key features of synchronous OLPE, including various coding Levels, category definitions, and observation tactics and procedures, as well as develop a systematic observation instrument for assessing online synchronous teaching in K-12 PE settings. We also focused on how virtual PE instruction incorporates family and community resources to support students’ learning.

METHOD

The following five phases were used to develop and establish the reliability of the System for Observing Virtual Real Time Lessons in Physical Education (SOVRTL-PE).

Specifically, the instrument was content validated, followed by Brewer and Jones (2002) and Russ et al. (2017), using an extensive literature search, observations of synchronous OLPE lessons, and consensus from a Delphi survey with several scholars/experts within the field (Boateng et al., 2018).

Phase □: Establishing on *A Priori* framework

An extensive literature review encompassing OLPE, CSPAP, and systematic observation in PE/physical activity was conducted to establish an *a priori* conceptual framework. For the initial framework, we conceptualized the quality of educational practices for synchronous OLPE, based on a range of the characteristics and recommendations/suggestions identified in the literature. Specifically, we identified seven major categories:

(a) Prerequisites/supportive learning environment: Teachers need to set criteria for skills tests and consider the use of online readiness tools (SHAPE America, 2018), check students' computer technology skills (Mohnsen, 2012; Williams et al., 2020), and provide technical support (Mohnsen, 2012)

(b) Curriculum: Teachers need to contribute to the goals of National PE standards and contents as well as influence the next generation of movers to become physically active and healthy for a lifetime (Daum & Buschner, 2012; SHAPE America, 2018)

(c) Student motivation: Teachers need to personalize student learning as a key component to online courses (SHAPE America, 2018), allow for student choice so students can choose where, when, and how to be physically active (Killian et al., 2020; SHAPE America, 2018; Williams et al., 2020) and use relevant technologies to enhance

motivation and accountability for OLPE lessons (Daum & Buschner, 2018; Kooiman & Sheehan, 2014)

(d) Student engagement: Teachers need to stimulate interaction with students, both between themselves and the student and between students to create a sense of community (i.e. video or chat message boards) (Daum & Buscher, 2012; NASPE, 2007; Nicholas & Ng, 2009; SHAPE America, 2018; Williams, 2013)

(e) Teacher feedback: Teachers need to deliver feedback immediately for learners to progress during motor skill instruction (Daum & Buschner, 2018) and provide timely and frequent feedback about student progress (International Society for K-12 Online and Blended Learning [iNACOL], 2011)

(f) Resource optimization: Teachers need to include synchronous activities such as audio or video conferencing, teacher presentations, the use of electronic whiteboards, shared forms, email, and chatting (Killian et al., 2020; Mohnsen, 2012; NASPE, 2007) in order to have students demonstrate, identify, and even correct a principle, strategy, and/or tactic (Daum, 2020) and integration of technology tools such as apps for smart devices and YouTube channels, which can provide multiple ways of learning and physical activity opportunity (Daum, 2020; Daum & Buschner, 2018)

(g) Assessment: Teachers need to ensure there is a wide range of formative and summative assessments throughout the entire course using different tools (e.g., video recording and online quiz or test; Daum, 2020; Daum & Buschner, 2018; SHAPE America, 2018)

Phase □: Expanding and refining the *A Priori* framework

Videotape analysis. After receiving approval from the Institutional Review Board at the University of South Carolina to conduct the study, the lead researcher observed and video recorded synchronous OLPE lessons in a semester-long secondary methods course that was converted from a face-to-face format to a fully online format (via Microsoft Teams) following the COVID-19 outbreak in the fall 2020 academic semester (15 weeks). The lead researcher observed a total of 20 OLPE lessons across 15 students (12 undergraduates and 3 graduates; 4 female and 11 male). Of these, 242 total minutes of recorded/videotaped (n=6) were collected, which ranged from 35 to 44 minutes (M = 40.3, SD = 2.94). Recorded videos included a wide range of information/examples of synchronous OLPE. Specifically, the researcher focused on objectives-based instruction aligned with PE state or national standards; instructional style; task presentation and other communication skills (e.g., set induction and lesson closure); management skills; content development; teacher feedback; motivational strategies; technology integration; leveraging home- and community-based resources to support students' learning; and assessment of students' learning (Benes & Alperin, 2019; Carson & Webster, 2020; Rink, 2020; SHAPE America, 2015, 2018). As supplementary material, the lead researcher also collected the three course instructors' observation field notes and feedback to the students (72 sets of notes). In addition, the lead researcher collected additional videos (n=15; a total 422 minutes) related to synchronous OLPE for K-12 from open resources (e.g., YouTube channel), which ranged from 24 to 41 minutes (M = 29.3, SD = 8.32).

As video data were collected, the lead researcher consistently reviewed/observed recorded videos to catalogue examples of synchronous OLPE. The *a priori* conceptual

framework guided initial observations, although the researcher also remained sensitive to unanticipated teaching behaviors/practices, lesson event, or situations in OLPE lessons. Video examples and an initial category system of OLPE were discussed with an academic advisor whenever unanticipated contexts or situations emerged. In such cases, if the identified behavior/situation was not readily catalogued using the *a priori* conceptual framework, the framework was modified/revised accordingly. This approach allowed for an in-depth analysis and understanding of the critical components/factors of synchronous OLPE practices. Consistent with established instrument development procedures, video viewings and discussions continued throughout video analysis and afterward to confirm and expand the category system (i.e., terms and definitions) until the observations yielded no additional insight (Russ et al., 2017; Thomas, Nelson, & Silverman, 2011; Weaver et al., 2014). Through this process, we developed 60 different variables (with definitions) and organized these across seven levels of focus for systematic observation.

Delphi survey. The Delphi method is a commonly used research methodology that utilizes several rounds of questioning to obtain anonymous consensus from content experts regarding the definition of an ambiguous term or idea (Avella, 2016; Hsu & Sandford, 2007). The purpose of this survey was to gain validate/consensus on the creation of appropriate variables and their definitions to help develop the instrumentation for assessing synchronous OLPE. Whereas the *a priori* framework and the recorded OLPE lessons were used to develop the initial definitions for each variable in the instrument, the results from the Delphi survey were subsequently used to confirm, refine, and/or validate the category system for SOVRTL-PE.

Participants in the survey were purposefully selected using an online search and references list from a literature analysis (i.e., Phase 1). Specifically, participants were selected based on whether they met at least one of the following criteria: (a) produced resources for OLPE, (b) are a university faculty member in a PE teacher education program, (c) have authorship in OLPE conceptual/theoretical, empirical, and/or professional literature and have published work as a first or corresponding author on virtual OLPE in the last 10 years with a focus on the role of technology, (d) have a record of developing systematic observation instruments in PE and/or physical activity, or (e) are a first author of a publication on CSPAP, particularly in the area of family and community engagement. The Delphi survey was accordingly sent electronically to 37 individuals via SurveyMonkey. An 80% level of agreement per each variable across participants was initially set to retain instrument definitions and obtain content validity (Boateng et al., 2018; Turoff & Linstone, 2002).

The survey was administered in two rounds. In the first round, participants reviewed, rated, and provided feedback on the draft of the SOVRTL-PE instrument. Questions for the first round were divided into two sections. Section 1 focused on the measurement variables and their definitions by each level, which included seven observation categories: lesson structure, lesson focus, lesson context, instructional style, teaching behavior, positive motivational aids, and technology integration. Section 2 focused on the recording technique for data collection. Participants were asked whether these categories should be included in the instrument and about the clarity of the instrument using agree, partially agree, or disagree as selection options. Open text fields were also included to allow for participant comments. Twenty-five out of thirty-seven

invited participants (68%) responded to the survey, which showed an overall agreement score as follows: agree (73.66%), partially agree (18.47%), and disagree (7.88%). After the first-round survey closed, the researchers used the participants' responses to revise the instrument. A total of 189 comments were received with the majority of the feedback related to making additions, modifications, and/or deletions to terms/definitions. Overall, a total of 30 changes were made. For instance, we made sub-codes within a task progression variable as a response to the request to add detail about the quality of the task presentation (Rink & Werner, 1989). In addition, we changed the name of the "technology glitch" variable to "technology malfunction" with a revised definition for clarification and also created a "technology challenge" variable. We did not modify any variables or definitions that were adopted verbatim from previously established instruments. For example, all of the variables at Level 2 (Lesson Focus) were directly derived from the System for Observing Fitness Instruction Time (SOFIT; McKenzie, 2002, 2015), which has been used extensively in PE research (McKenzie & Smith, 2017).

Subsequently, the lead researcher invited the Delphi participants to engage in a second round of the survey, which followed the same procedures and tasks as the first round. Eighteen individuals out of 25 (70%) responded to the second round of the survey, which received an overall score across all variables as follows: agree (88.62%), partially agree (7.30%), and disagree (4.08%). The second round yielded no further insights; therefore, no further rounds were pursued.

Phase □: Devising a system for coding

The first and second author experimented with multiple recording techniques to code the variables in SOVRTL-PE by watching videos of synchronous OLPE lessons and

attempting to reliably capture what they observed. A partial interval recording observation system was ultimately selected for SOVRTL-PE to allow for the assessment of multiple events simultaneously (Hintze, Volpe, & Shapiro, 2002; McKenzie & van der Mars, 2015). Additionally, the decision was made to recommend that SOVRTL-PE be used primarily as an analysis instrument for video recorded data. Video recordings have an advantage in that they permit the observer to review the teaching behaviors and lesson context related to variables as many times as needed to make coding decisions, whereas live coding/observation requires the observer to code in real time (Hintze et al., 2002). It should potentially help to reduce inaccurate coding and/or missing data more than live conditions (McKenzie & van der Mars, 2015). The following coding procedures were created: (1) Begin observing at the start of the recorded lesson and continue until the recording ends (this permits the observer to understand the lesson context, which is important to identifying the appropriate codes during interval recording); (2) Observe for 10 seconds, then pause the video and record the appropriate codes, starting at Level 1 and continuing through Level 7 in sequence, for as long as needed. Resume the video for another 10 seconds, then pause to record again. Continue for the duration of the lesson; and (3) For each interval, choose all relevant codes for Level 1-7 (i.e., code all variables for which there is evidence).

The observer must also follow several ground rules when coding to ensure reliable use of the instrument. First, at Level 5, the observer should code "Monitoring" or "Other" if the teacher's camera if off, code "Other" for when there is a technology malfunction or technology challenge at Level 3, and code "Monitoring" in all other cases. Second, the observer should change the code if evidence emerges in a subsequent interval

indicating a change is needed to a previous code. Third, the observer should code "Other" when the teacher or a student says or does something that is not clear.

Phase IV: Pilot testing

Pilot testing of SOVRTL-PE was conducted with six graduate students enrolled in a 3-week summer course on observational methods (mostly systematic observation) in PE and physical activity instructional contexts. The students were already trained in the knowledge and concepts of systematic observation and had been exposed to various established systematic observation tools in the field. The first and second authors presented the new tool to the class and answered questions about the variables, their definitions, the recording procedures, and the ground rules. Students were shown multiple video clips from synchronous OLPE lessons and asked to independently code what they observed using the tool. After coding each clip, the researchers led discussions to resolve coding discrepancies. After approximately 75 minutes of training, the students engaged in a reliability testing exercise, which involved independently coding the first several minutes of a recorded lesson they had not seen before. The first author had previously coded the same lesson to establish a criterion code for establishing inter-rater reliability. This testing exercise resulted in high agreement percentages with the criterion codes for each graduate student (.89-.97). There were some minor modifications made as a result of the pilot test. For instance, we changed the name of Level 2 from “Lesson Context” to “Lesson Event” to better conceptually support selecting multiple codes at this level and revised the definition of the “Assessment” variable within Level 5 (Teaching Behavior) to more clearly discriminate between assessment behavior and “Monitoring”, which is

another variable at the same level. The final variables included in each level and their definitions are presented in Table 4.1.

Phase □: Formal reliability testing

Interobserver reliability. Established recommendations for testing reliability were followed by a specific sequence of steps: (a) orientation to the SOVRTL-PE instrument, (b) committing each category/code to memory, (c) video practice with discussions, and (d) reliability testing (Brewer & Jones, 2002; McKenzie & van der Mars, 2015; Pope et al., 2002; Russ et al., 2017; van der Mars, 1989). Two PE doctoral students not involved in developing the instrument participated in a training that involved the above-mentioned steps and were then tested on their ability to achieve interobserver reliability using the instrument. Reliability testing was divided into two phases designed to progress trainees from simpler to more complex coding requirements. Phase 1 was conducted as a preliminary test using multiple video clips (~5–10 seconds), each focusing on a different level of the tool. This phase was intended to help observers focus on one level at a time and better understand the variables at each level. In Phase 2, trainees observed and coded a full length, video recorded synchronous OLPE lesson (~40 minutes), which they had not previously seen. The first author also coded the lesson to establish criterion codes for determining interobserver reliability (Webster et al., 2013). Interobserver reliability (IOR) was calculated as an interval-by-interval agreement for each level using the formula presented below (Moon et al., 2020; Russ et al., 2017):

$$\frac{\text{Number of Agreements}}{\text{Number of agreements} + \text{Disagreements}} \times 100 = \%\ \text{Agreement}$$

Above 80% agreement (.91 and .94) was reached against a criterion score by both trainees, which was deemed acceptable (Table 4.2.; Brewer & Jones, 2002; van der Mars, 1989). Approximately 3.5 hours of training across two days were required to achieve this reliability standard.

Intraobserver reliability. Intraobserver reliability is a measure of an instrument's ability to be used in the same way by the same person over time (Thomas et al., 2011). After a 10-day period, the two trainees re-coded the same lesson (Webster et al., 2013) and the same formula used to calculate interobserver reliability was used to calculate intraobserver reliability. A high level of overall agreement was reached (.95).

DISCUSSION

To our knowledge, SOVRTL-PE is the first systematic observation instrument for comprehensively measuring synchronous OLPE in K-12 settings. This instrument addresses the need for objective measurements of synchronous OLPE instruction, which also includes how virtual PE practices can optimize family and community resources within the CSPAP framework to support students' learning (CDC, 2019; Institute of Medicine, 2013). Overall, the result was found to follow accepted levels of inter-observer and intra-observer values. SOVRTL-PE could be used by trained observers in future studies of synchronous OLPE instruction or in the measurement and analysis of OLPE teaching practices. However, there is a need to assess the instrument in a variety of video conferencing platforms (e.g., Webex, Blackboard Collaborate) and with additional observers to further determine its generalizability and to identify more precisely the amount of time for quality observer training, which affects the reliable and accurate use of the system (McKenzie & van der Mars, 2015).

While SOVRTL-PE was found to be valid and reliable overall through multiple phases, some variables were not observed enough to establish content validity. Specifically, of the 62 variables, the variables “Desist”, “Polling”, “Closed Captioning” were present less frequently than others. The use of established literature, recorded video, and the Delphi survey to derive SOVRTL-PE variables suggests that while these variables might not be as commonly reflected in teachers’ OLPE practices than other variables included in the instrument, they still should contribute to an understanding the features of synchronous OLPE lessons and teaching practices. For instance, the “Desist” variable (within Level 6) was recommended by some of the Delphi experts in this study and has been explored as an established definition of instructional strategy to prevent students’ misbehavior (Kounin, 1970). Further research with this instrument is needed to examine these variables and can help to establish appropriate instructional standards for synchronous OLPE lessons.

This study has several limitations. Even though we used multiple examples of synchronous OLPE recorded video, the relatively small amount of video minutes (i.e., a total 664 minutes) in Phase 2 of the study limits the generalizability of the results. While contextual variables (within Level 3) were considered in selecting recorded videos for observation, it is possible that the SOVRTL-PE category system does not fully represent the range of synchronous OLPE practices/instructions that would be found across different lessons and settings. In addition, the SOVRTL-PE was not validated in live coding conditions (only observed through recorded video). Thus, a follow-up study is warranted to test the plausibility of live coding. Finally, while the SOVRTL-PE captures different types of OLPE teaching practices within a variety of both preservice and in-

service teaching contexts, it does not provide detailed information about specific tasks/activities or students' responses during OLPE lessons. Researchers should consider collecting additional evidence/information through other resources (e.g., field notes) when using SOVRTL-PE or should modify the instrument to increase its sensitivity to such information.

Despite the limitations discussed above, this study allows researchers, educators, and practitioners to gain insight into the key features of synchronous OLPE and provides the SOVRTL-PE instrument for documenting teachers' practices in K-12 settings. For instance, SOVRTL-PE will enable the development of a descriptive research base for synchronous OLPE for K-12 lessons and will enhance measurement of OLPE-based interventions irrespective of the specific curriculum. Teacher educators can also use the instrument to help develop in-service or preservice teachers' teaching practices for quality of synchronous OLPE lessons, which can be better optimized with family and community resources to support students' learning.

CONCLUSION

This study has shown that the SOVRTL-PE is a valid observation system for recording lesson structure, lesson focus, lesson event, instructional style, teaching behaviors, motivational aids, and technology integration during synchronous OLPE lessons. In this sense, SOVRTL-PE will provide researchers and practitioners with a valid instrument to objectively assess/measure and quantify synchronous OLPE practices. While continued research is needed to establish the reliability and validity of SOVRTL-PE in a broad array of virtual teaching contexts, such as in live conditions or through various video conferencing platforms, SOVRTL-PE uniquely generates information

regarding how to optimize family and community resources and how to align the CSPAP framework with OLPE. Increased study in the area of OLPE will be foundational to informing measures related to the use of quality synchronous OLPE in both preservice teacher education and K–12 PE. Ultimately, future research using SOVRTL-PE can help identify best practices for synchronous OLPE and establish recommendations for synchronous OLPE programming.

Table 4.1 Operational Definitions Scores for SOVRTL-PE Variables

SOVRTL-PE Variables	Definitions
Level 1: Lesson Structure	
Whole Group (WG)	The class is meeting as one group.
Breakout (B)	The class is divided into breakout rooms.
Level 2: Lesson Focus (adapted from McKenzie, 2015)	
General Content (GC)	Lesson time when students are not intended to be involved in physical education content, including transition, management, and break times. Transition includes time allocated to managerial and organizational activities related to instruction such as team selection, changing equipment, moving from one space to another, changing stations, teacher explanation of organizational arrangement, and changing activities within a lesson. Management includes time devoted to class business that is unrelated to instructional activity such as taking attendance, discussing a field trip, or collecting money for class pictures. Break includes time devoted to rest and/or discussion of non-subject matter topics such as getting a drink of water, talking about last night's ball game, telling jokes, celebrating the birthday of a class member, or discussing the results of a class election.
Knowledge Content (KC)	Lesson time when the primary focus is on student acquisition of knowledge related to physical education, not being physically active. Knowledge is typically related to: (a) Physical activity and fitness (i.e., information related to physical activity or physical fitness concepts, including endurance, strength, and flexibility), and (b) General Knowledge (information related to areas other than physical activity and fitness, such as history, technique, strategy, rules, and social behavior).

Fitness (F)	Lesson time allocated to activities whose major purpose is to alter the physical state of the individual in terms of cardiovascular endurance, strength, or flexibility. This includes aerobic dance, calisthenics, distance running, weight training, agility training, fitness testing, and warm-up and cool down activities. Code relays conducted with more than three per team as games, not fitness.
Game Play (G)	Activity time devoted to the application of skills in a game or competitive setting. Game participants generally perform without major intervention from the instructor, such as during volleyball and tag games, balance beam routines, and folk dance performances.
Free Play (FP)	Free play time during which physical education instruction is not intended. This time resembles recess during which students may select to participate or not.
Other (O)	The lesson focus does not fit any of the other codes.

Level 3: Lesson Event	
Set Induction (S)	The teacher is focusing on instructional content at the start of the lesson before transitioning to specific learning tasks.
Management (MA)	The teacher is organizing logistical details.
Task Presentation (TP)	The teacher is presenting a specific task, activity, or other content that students will try. *Sub-codes: 1) Informing task - The first task in a progression of content. 2) Refinement task - A teacher move that communicates a concern for the quality of student performance (e.g., work to get your toss a little higher). 3) Extension task - A teacher move that communicates a concern for changing the complexity or difficulty of student performance. 4) Application task - A teacher move that communicates a concern for moving the student focus from how to do the movement to how to use the movement.
Feedback (F)	The teacher is offering a comment about student performance.
Student Practice without PA (SPW)	The students are doing tasks/activities assigned by the teacher that do not involve physical activity (e.g., a written quiz, a task sheet).

Student Practice with PA (SPA)	The students are doing tasks/activities assigned by the teacher that involve physical activity.
Closure (C)	The teacher is ending the lesson.
Technology Malfunction (TM)	Something goes wrong with technology being used for the lesson (e.g., mic does not work, Internet connection is lost).
Technology Challenge (TC)	The teacher is struggling to use functioning technology available for the lesson.
Other (O)	The lesson context does not fit any of the other codes.
Level 4: Instructional Style	
Lecturing/Directing (L/D)	The teacher is directing the lesson without asking the students any questions.
Questioning (Q)	The teacher is asking the students questions.
Questioning with Wait Time (QW)	The teacher is asking students questions and waiting, as needed, for responses.
Troubleshooting (T)	The teacher is working on a technology issue.
Level 5: Teaching Behavior	
Objectives (OB)	The teacher is addressing the lesson's learning objectives (i.e., learning outcomes resulting from participation in the lesson).
Agenda (A)	The teacher is providing an agenda for the lesson (i.e., what activities the class will do throughout the lesson)
Student Technology Support (STS)	The teacher is addressing student technology needs.
Rules (RU)	The teacher is addressing the class rules or expectations for student conduct.
Managerial Directions (MD)	The teacher is giving students directions to organize a task/activity.
Consequence (CQ)	The teacher is imposing a consequence for a student breaking a rule.

Demonstration (D)	The teacher is showing/modeling for students how to do a task/activity (psychomotor, cognitive, or affective) before having the students do the task/activity. * Sub-codes: 1.With Content Enrichment; 2. Full; 3. In Context.
Content Enrichment (CE)	The teacher is explaining, summarizing, or defining the critical elements/information for a task/activity without a demonstration.
Specific Congruent Feedback (SC)	The teacher is providing detailed information when giving feedback that is aligned with the focus of a previous task presentation during the lesson.
General Feedback (GF)	The teacher is providing feedback that lacks detailed information.
Participation (PR)	The teacher is participating in the activity that students are doing.
PE Promotion Outside Class (PH)	The teacher is explaining how students can extend their physical education learning outside of class.
PA Promotion Outside Class (PA)	The teacher is encouraging students to be physically active outside of class.
Assessment (A)	The teacher is collecting data about student learning, performance, or behaviors.
Monitoring (MO)	The teacher is silently observing, supervising, or listening to students during periods of instruction or practice.
Other (O)	The teacher behavior does not fit any of the other codes.
Level 6: Motivational Aids	
Choice (CH)	The teacher is giving students options.
Personalization (P)	The teacher is casually conversing with students, sharing personal stories, obtaining personal information from students, or relating the lesson to other things the students might know or care about.
Encouragement (E)	The teacher is giving students encouragement about their learning.
Reinforcement (R)	The teacher is making a statement that reinforces student behavior as it relates to classroom management (e.g., "Thank you for putting your comment in the chat box", "I appreciate

	those of you who remembered to keep your mic turned off during others' presentations").
Incentive (IN)	The teacher is providing students with a reward (e.g., making a student MVP, providing extra free time at the end of lesson).
Desist (DT)	The teacher is taking actions to stop off-task or disruptive student behavior.
None (N)	No motivational aids are evident.
Level 7: Technology Integration	
Chat Box (CB)	The teacher or students are using the chat function.
Closed Captioning (CC)	The teacher is using the closed captioning function.
Presentation Slides (PS)	The teacher or students are sharing a slide deck (e.g., PowerPoint, Prezi).
Smart Device (SD)	The teacher or students are using a smart phone, smart watch, tablet, or other smart device (in addition to the device they used to join the meeting).
Shared Form (SF)	The teacher or students are working on a document that can be shared/edited in real time (e.g., Google Sheet).
Camera (CM)	The teacher's webcam is on. * Sub-codes: 1.Teacher Visible (TV); 2. Teacher Not Visible (TN).
Pin/Spotlight (P)	The teacher or students are using the pin/spotlight function.
Mic (MI)	The teacher or students use a microphone (built into computer or otherwise).
Video (V)	The teacher or students are using or creating a video.
Email (EM)	The teacher or students are using email.
Screenshare (SH)	The teacher or students are sharing content from their screen.
Music (MS)	The teacher or students are playing music audibly.
Sound Effects (SE)	The teacher or students are using digital sound effects.
Whiteboard (WH)	The teacher or students use digital whiteboard technology.
Polling (PL)	The teacher or students are using the polling function.

Emoji (EM)	The teacher or students are using an emoji (e.g., raised hand, smiley face).
Timer (TM)	A timer is being used (e.g., as part of a video, in a slideshow).
Supplementary Device (SU)	The teacher is using more than one device to log in to the lesson.
Technology Supplement (TS)	The teacher is using a non-technological aid to teach.
Other (O)	The teacher or students are using a technology tool that does not fit any of the other codes.

Table 4.2 Inter-rater Reliability Scores for SOVRTL-PE Variables

		Level 1 (%)	Level 2 (%)	Level 3 (%)	Level 4 (%)	Level 5 (%)	Level 6 (%)	Level 7 (%)	Total agreement (%)
Phase 1	Graduate student 1	100	N/A	91.55	90.37	93.21	90.17	94.65	N/A
	Graduate student 2	100	N/A	90.00	88.75	92.78	88.29	92.23	N/A
Phase 2	Graduate student 1	100	N/A	93.01	93.78	92.27	89.75	95.88	94.12
	Graduate student 2	100	N/A	89.17	91.13	87.88	87.05	93.62	91.48

School: ________________ Teacher name: __________ Teacher location: _______ Grade:______________

Number of students: ____ Lesson length: _____(min) Observation Date: ______ Observer: _____________

Figure 4.1 SOVRTL-PE Coding Sheet.

CHAPTER 5

DISCUSSION

The overarching goal of this dissertation was to examine comprehensive school physical activity program (CSPAP)-aligned strategies, particularly beyond traditional physical education (PE), for enhancing students' PE learning. Although each of the five components of a CSPAP contributes uniquely to facilitate children's physical activity (PA), little is known about the extent to which the different components of a CSPAP can support students' achievement of the SHAPE National Standards. This dissertation consists of three studies that in tandem contribute to the CSPAP literature by exploring the potential of CSPAPs, whether facilitated in-person or remotely, for supporting students' PE learning.

Study 1 was a systematic review and meta-analysis of the effectiveness of CSPAP-aligned PA interventions on the development of students' motor competence (MC). Study 2 was a systematic review and meta-analysis of the effectiveness of CSPAP-aligned PA interventions on increasing students' social and emotional learning (SEL). The results of both studies reveal that the PA intervention programs had a positive effect on the development of MC and SEL. The main findings from both Studies 1 and 2 indicate that the PE component, as either a single component intervention approach or as part of a multicomponent approach, is a foundational component to improving students' MC and SEL. In addition, both studies highlight that ongoing teacher training and

continuous support during implementation are critical elements for successfully achieving the desired outcomes.

Another notable finding from both studies is that multiple component approaches appear to be promising avenues in the development of students' MC and SEL. However, there is insufficient evidence to reach conclusions about the effectiveness of multicomponent approaches to improve these outcomes, and more research is warranted to explore the synergistic potential of combining different CSPAP components to enhance children's PE learning. Additionally, further investigation is needed to determine which strategies are most effective for optimizing family and community engagement and staff involvement components of the CSPAP framework to support the development of MC and/or SEL in school-aged children.

Based on the implications from both Studies 1 and 2, this dissertation generated evidence to support the continued pursuit of CSPAP-aligned implementation programs for developing elementary children's MC and SEL. In particular, targeting the quality of such programs, as well as measuring fidelity of implementation through developing component-specific objective measures, will be efficacious strategies that could help to increase program effectiveness. Following these recommendations will be critical to understanding the linkage between program implementation, uptake, and effectiveness, and will provide formative information for monitoring and adjusting CSPAP initiatives to better meet desired learning outcomes in schools.

Study 3 helps to address the question "What's going on in synchronous online physical education (OLPE) practices?" Specifically, the study illustrates an overview of key characteristics of synchronous OLPE and introduces the System for Observing

Virtual Real Time Lessons in Physical Education (SOVRTL-PE) instrument for systematically assessing synchronous OLPE practices in K-12 PE. The study also underscores the importance of incorporating family and community resources to better understand how to support students' learning during virtual PE instruction.

Globally, a growing body of literature related to online learning has been emerging in recent years. Study 3 provides one avenue for assessing and improving professional practices amid increased shifts toward virtual teaching and learning alternatives in PE. SOVRTL-PE affords exploration into the quality of online teaching practices and will help researchers and teacher educators conceptualize the integration of evidence-based synchronous OLPE practices in tandem with family and community resources. Additionally, SOVRTL-PE can serve to support programs' goals and objectives in teacher education programs and professional development programming related to synchronous OLPE teaching. Overall, the SOVRTL-PE instrument has the potential to expand resources for objectively measuring or assessing synchronous OLPE practices. Utilizing this instrument for examining and developing virtual teaching practices within the CSPAP framework can allow practitioners to base policy and practice decisions on objective measurement data.

Is it possible for CSPAPs to make their way into routine school-based practice and effect meaningful change in students' development and the trajectory of their PA behavior? As the application of the CSPAP framework moves into its next iteration of research and practice, continued focus on maximizing the potential of opportunities provided in each CSPAP component is imperative to increasing the uptake and success of this multicomponent approaches to youth PA promotion. This dissertation contributes to

the understanding and facilitation of effective CSPAP-aligned approaches to achieve student outcomes identified as important in the PE profession. Future CSPAP research should continue to explore how to evoke positive changes in the development of students' PE learning. In particular, researchers should strive to discover what types of CSPAP-aligned PA experience best leverage students' knowledge and skills in line with SHAPE's National Standards. This will provide evidence for the mechanisms through which PA interventions are optimally effective and will help establish recommendations and policies for school and community practices that might finally make a significant impact in efforts to promote youth PA, and sufficiently move the needle toward desired health-related outcomes in school-aged children and adolescents.

REFERENCES

Ajzen, I. (1991). The theory of planned behavior. *Organizational behavior and human decision processes, 50*(2), 179–211.

Allen, I. E., & Seaman, J. (2016). *Online report card: Tracking online education in the United States*. Babson Survey Research Group. Babson College, 231 Forest Street, Babson Park, MA 02457.

Andrade, A., Cruz, W. M. D., Correia, C. K., Santos, A. L. G., & Bevilacqua, G. G. (2020). Effect of practice exergames on the mood states and self-esteem of elementary school boys and girls during physical education classes: A cluster-randomized controlled natural experiment. *Plos one*, *15*(6), e0232392.

Ang, S., & Penney, D. (2013). Promoting social and emotional learning outcomes in physical education: Insights from a school-based research project in Singapore. *Asia-Pacific Journal of Health, Sport and Physical Education, 4*(3), 267–286.

Annesi, J. J., Smith, A. E., Walsh, S. M., Mareno, N., & Smith, K. R. (2016a). Effects of an after-school care-administered physical activity and nutrition protocol on body mass index, fitness levels, and targeted psychological factors in 5-to 8-year-olds. *Translational behavioral medicine*, *6*(3), 347–357.

Annesi, J. J., Walsh, S. M., & Greenwood, B. L. (2016b). Increasing children’s voluntary physical activity outside of school hours through targeting social cognitive theory variables. *Journal of primary care & community health*, *7*(4), 234–241.

Annesi, J. J., Walsh, S. M., Greenwood, B. L., Mareno, N., & Unruh□Rewkowski, J. L. (2017). Effects of the Youth Fit 4 Life physical activity/nutrition protocol on body mass index, fitness and targeted social cognitive theory variables in 9□to 12□year□olds during after□school care. *Journal of paediatrics and child health*, *53*(4), 365–373.

Armour, K., Quennerstedt, M., Chambers, F., & Makopoulou, K. (2017). What is ‘effective’CPD for contemporary physical education teachers? A Deweyan framework. *Sport, education and society*, *22*(7), 799–811.

Ames, C., & Archer, J. (1988). Achievement goals in the classroom: Students' learning strategies and motivation processes. *Journal of educational psychology*, *80*(3), 260–267.

Avella, J. R. (2016). Delphi panels: Research design, procedures, advantages, and challenges. *International Journal of Doctoral Studies*, *11*(1), 305–321.

Banich, M. T. (2009). Executive function: the search for an integrated account. *Current Directions in Psychological Science, 18*(2), 89–94.

Barcelona, J., Centeio, E. E., & Parker, M. (2020). Integrating health and physical activity with social emotional learning in before-and after-school programs. In *Before-and After-School Physical Activity Programs* (pp. 57-68). Routledge.

Barnett, L. M., Lai, S. K., Veldman, S. L. C., Hardy, L. L., Cliff, D. P., Morgan, P. J., . . . Okely, A. D. (2016). Correlates of gross motor competence in children and adolescents: A systematic review and meta-analysis. *Sports Medicine, 46*(11), 1663–1688.

Barnett, L. M., Stodden, D., Cohen, K. E., Smith, J. J., Lubans, D. R., Lenoir, M., ... & Morgan, P. J. (2016). Fundamental movement skills: An important focus. *Journal of Teaching in Physical Education*, *35*(3), 219–225.

Barnett, L. M., Vazou, S., Abbott, G., Bowe, S. J., Robinson, L. E., Ridgers, N. D., & Salmon, J. (2016). Construct validity of the pictorial scale of perceived movement skill competence. *Psychology of Sport and Exercise, 22*, 294–302.

Barnett, L. M., van Beurden, E., Morgan, P., Brooks, L., & Beard, J. (2009). Childhood motor skill proficiency as a predictor of adolescent physical activity. *Journal of Adolescent Health. 44(3)*, 252–259.

Barry, M. M., Clarke, A. M., & Dowling, K. (2017). Promoting social and emotional well-being in schools. *Health Education, 117*(5), 434–451.

Bartholomew, J. B., & Jowers, E. M. (2011). Physically active academic lessons in elementary children. *Preventive Medicine, 52*(1), S51–54.

Benes, S. S., & Alperin, H. L. (2019). *Lesson planning for skills-based health education: Meeting secondary-level national standards*. Champaign, IL: Human Kinetics.

Best, J. R. (2010). Effects of physical activity on children's executive function: Contributions of experimental research on aerobic exercise. *Developmental Review, 30*(4), 331–351.

Biddle, S., Ciaccioni, S., Thomas, G., & Vergeer, I. (2019). Physical activity and mental health in children and adolescents: An updated review of reviews and an analysis of causality. *Psychology of Sport & Exercise*, *42*, 146–155.

Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quiñonez, H. R., & Young, S. L. (2018). Best practices for developing and validating scales for health, social, and behavioral research: a primer. *Frontiers in public health, 6*, 149.

Bolger, L. E., Bolger, L. A., O'Neill, C., Coughlan, E., O'Brien, W., Lacey, S., & Burns, C. (2019). The Effectiveness of Two Interventions on Fundamental Movement Skill Proficiency among a Cohort of Irish Primary School Children. *Journal of Motor Learning and Development*, *7*(2), 153–179.

Brewer, C. J., & Jones, R. L. (2002). A five-stage process for establishing contextually valid systematic observation instruments: The case of rugby union. *The Sport Psychologist*, *16*(2), 138–159.

Brian, A., Getchell, N., True, L., De Meester, A., & Stodden, D. F. (2020). Reconceptualizing and operationalizing Seefeldt's proficiency barrier: Applications and future directions. *Sports Medicine*, 1–12.

Brian, A., Goodway, J. D., Logan, J. A., & Sutherland, S. (2017). SKIPing with Head Start teachers: Influence of T-SKIP on object-control skills. *Research Quarterly for Exercise and Sport*, *88*(4), 479–491.

Bridgeland, J., Bruce, M., & Hariharan, A. (2013). *The Missing Piece: A National Teacher Survey on How Social and Emotional Learning Can Empower Children and Transform Schools*. Washington, DC: Civic Enterprises/Hart Research Associates.

Bronfenbrenner, U. (1986). Ecology of the family as a context for human development: Research perspectives. *Developmental psychology*, *22*(6), 723.

Brown, T. C., & Fry, M. D. (2014). Evaluating the pilot of strong girls: A life skills/physical activity program for third and fourth grade girls. *Journal of Applied Sport Psychology*, *26*(1), 52–65.

Buišić, S., & Đorđić, V. (2019). The effectiveness of Hellison's model of personal and social responsibility in physical education teaching. *Journal of Facta Universitatis, Series: Physical Education and Sport*, 663–675.

Bundy, A., Engelen, L., Wyver, S., Tranter, P., Ragen, J., Bauman, A., ... & Naughton, G. (2017). Sydney playground project: a cluster□randomized trial to increase physical activity, play, and social skills. *Journal of school health*, *87*(10), 751–759.

Burns, R. D., Fu, Y., Fang, Y., Hannon, J. C., & Brusseau, T. A. (2017a). Effect of a 12-week physical activity program on gross motor skills in children. *Perceptual and Motor Skills*, *124*(6), 1121–1133.

Burns, R. D., Fu, Y., Hannon, J. C., & Brusseau, T. A. (2017b). School physical activity programming and gross motor skills in children. *American journal of health behavior*, *41*(5), 591–598.

Burrows, E. J., Keats, M. R., & Kolen, A. M. (2014). Contributions of after school programs to the development of fundamental movement skills in children. *International journal of exercise science*, *7*(3), 236–249.

Buschner, C. (2006). Online physical education: Wires and lights in a box. *Journal of Physical Education, Recreation & Dance, 77*(2), 3–8.

Carlson, J. A., Engelberg, J. K., Cain, K. L., Conway, T. L., Geremia, C., Bonilla, E., ... & Sallis, J. F. (2017). Contextual factors related to implementation of classroom physical activity breaks. *Translational Behavioral Medicine, 7*(3), 581–592.

Carson, R. L., Castelli, D. M., Beighle, A., & Erwin, H. (2014). School-based physical activity promotion: A conceptual framework for research and practice. *Childhood Obesity*, *10*(2), 100–106.

Carson, R. L., & Webster, C. A. (Eds.) (2020). *Comprehensive school physical activity programs: Putting research into evidence-based practice*. Champaign, IL: Human Kinetics.

Cecchini, J. A., Montero, J., Alonso, A., Izquierdo, M., & Contreras, O. (2007). Effects of personal and social responsibility on fair play in sports and self-control in school-aged youths. *European Journal of sport science, 7*(4), 203–211.

Centers for Disease Control and Prevention. (2015). *National framework for physical activity and physical education*. Centers for Disease Control and Prevention, US Department of Health and Human Services. https://www.cdc.gov/healthyschools/physicalactivity/pdf/National_Framework_Physical_Activity_and_Physical_Education_Resources_Support_CSPAP_508_tagged.pdf

Centers for Disease Control and Prevention [CDC]. (2010). *The Association between School Based Physical Activity, Including Physical Education, and Academic Performance.* Atlanta, GA: U.S. Department of Health and Human Services.

Centers for Disease Control and Prevention. (2019). *Increasing physical education and physical activity: A framework for schools*. Atlanta, GA: US Department of Health and Human Services.

Chan, C., Ha, A., & Ng, J. Y. Y. (2016). Improving fundamental movement skills in Hong Kong students through an assessment for learning intervention that emphasizes fun, mastery, and support: the A + FMS randomized controlled trial study protocol. *Springer Plus, 5*, 724.

Chan, C. H., Ha, A. S., Ng, J. Y., & Lubans, D. R. (2019). The A+ FMS cluster randomized controlled trial: An assessment-based intervention on fundamental movement skills and psychosocial outcomes in primary schoolchildren. *Journal of science and medicine in sport*, *22*(8), 935–940.

Chen, L. (2016). Impacts of flipped classroom in high school health education. *Journal of Educational Technology Systems, 44*(4), 411–420.

Chen, W., Mason, S., Hypnar, A., & Bennett, A. (2016). Assessing motor skill competency in elementary school students: A three-year study. *Journal of sports science & medicine*, *15*(1), 102–110.

Ciotto, C. M., & Gagnon, A. G. (2018). Promoting social and emotional learning in physical education. *Journal of Physical Education, Recreation & Dance*, *89*(4), 27–33.

Cipriani, K., Richardson, C., & Roberts, G. (2012). Family and community involvement in the comprehensive school physical activity program. *Journal of Physical Education, Recreation & Dance*, *83*(7), 20–26.

Clark, J. E. (1994). Motor development. In V. S. Ramachandran (Ed.), *Encyclopedia of Human Behavior* (Vol. 3, pp.163-190). San Diego, CA: Academic Press.

Cliff, D. P., Wilson, A., Okely, A. D., Mickle, K. J., & Steele, J. R. (2007). Feasibility of SHARK: A physical activity skill-development program for overweight and obese children. *Journal of Science and Medicine in Sport*, *10*(4), 263–267.

Cocca, A., Espino Verdugo, F., Ródenas Cuenca, L. T., & Cocca, M. (2020). Effect of a Game-Based Physical Education Program on Physical Fitness and Mental Health in Elementary School Children. *International Journal of Environmental Research and Public Health*, *17*(13), 4883.

Cohen, K. E., Morgan, P. J., Plotnikoff, R. C., Callister, R., & Lubans, D. R. (2015). Physical activity and skills intervention: SCORES cluster randomized controlled trial. *Medicine and science in sports and exercise*, *47*(4), 765–774.

Collaborative for Academic, Social, and Emotional Learning. (2020). *Core SEL competencies*. https://casel.org/

Collins, H., Booth, J. N., Duncan, A., & Fawkner, S. (2019). The effect of resistance training interventions on fundamental movement skills in youth: a meta-analysis. *Sports medicine-open*, *5*(1), 1–16.

Colvin, A.Y., Markos, N.J.E., & Walker, P.J. (2016). *Teaching fundamental motor skills, SHAPE America set the Standard* (3rd ed.). Champaign, IL: Human Kinetics,

Coolkens, R., Ward, P., Seghers, J., & Iserbyt, P. (2018). The effect of organized versus supervised recess on elementary school children's participation, physical activity, play, and social behavior: A cluster randomized controlled trial. *Journal of physical activity and health*, *15*(10), 747–754.

Corcoran, R. P., Cheung, A. C. K., Kim, E., & Xie, C. (2018). Effective universal school-based social and emotional learning programs for improving academic achievement: A systematic review and meta-analysis of 50 years of research. *Educational Research Review, 25*, 56–72.

Currie, D. (2016). *Growing up unequal: gender and socioeconomic differences in young people's health and well-being. Health Behaviour in School-aged Children (HBSC) study: international report from the 2013/2014 survey* (No. 7). World Health Organization.

Daum, D. N. (2020). Thinking about Hybrid or Online Learning in Physical Education? Start Here! Editor: Brian Mosier. *Journal of Physical Education, Recreation & Dance, 91*(1), 42–44.

Daum, D. N., & Buschner, C. (2018). Research on teaching blended and online physical education. In R. Ferdig & K. Kennedy (Eds.), *Handbook of research on K-12 online and blended learning* (2nd ed., pp. 321–334). ETC Press. https://figshare.com/articles/Handbook_of_Research_on_K12_Online_and_Blended_Learning_Second_Edition_/668681

Daum, D. N., & Buschner, C. (2012). The status of high school online physical education in the United States. *Journal of Teaching in Physical Education, 31*(1), 86–100.

Daum, D. N., & Woods, A. M. (2015). Physical education teacher educator's perceptions toward understanding of K-12 online physical education. *Journal of Teaching in Physical Education, 34*(4), 716–724. https://doi.org/10.1123/jtpe.2014-0146

Deci, E. L., & Ryan, R. M. (2008). Self-determination theory: A macrotheory of human motivation, development, and health. *Canadian psychology/Psychologie canadienne*, *49*(3), 182.

Deci, E. L., & Ryan, R. M. (2012). *"Self-determination theory," in Handbook of Theories of Social Psychology*, eds P. A. M. Van Lange, A. W. Kruglanski, and E. T. Higgins (Thousand Oaks, CA: Sage), 416–437.

Deeks, J. J., Higgins, J. P., Altman, D. G., & Cochrane Statistical Methods Group. (2019). Analysing data and undertaking meta□analyses. *Cochrane handbook for systematic reviews of interventions*, 241–284.

D’Hondt, E., Deforche, B., De Bourdeaudhuij, I., & Lenoir, M. (2009). Relationship between motor skill and body mass index in 5-to 10-year-old children. *Adapted Physical Activity Quarterly, 26*(1), 21–37.

Diamond, A. (2000). Close interrelation of motor development and cognitive development and of the cerebellum and prefrontal cortex. *Child Development, 71*, 44–56.

Diamond, A. (2013). Executive functions. *Annual Review of Psychology, 64*, 135–168.

Diamond, A., & Lee, K. (2011). Interventions shown to aid executive function development in children 4 to 12 years old. *Science, 333*(6045), 959–964.

Diamond, A., & Ling, D. S. (2015). Conclusions about interventions, programs, and approaches for improving executive functions that appear justified and those that, despite much hype, do not. *Developmental Cognitive Neuroscience*, *18*, 34–48.

Digital Learning Collaborative. (2019). Snapshot 2019: A review of K-12 online, blended, and digital learning. https://www.digitallearningcollab.com/

Digital Learning Collaborative. (2020). Snapshot 2020: A review of K-12 online, blended, and digital learning. https://www.digitallearningcollab.com/snapshot-pubs

Dinkel, D., Schaffer, C., Snyder, K., & Lee, J. M. (2017). They just need to move: Teachers' perception of classroom physical activity breaks. *Teaching and Teacher Education, 63*, 186–195.

Donnelly, J. E., & Lambourne, K. (2011). Classroom-based physical activity, cognition, and academic achievement. *Preventive Medicine, 52*(1), S36eS42.

Domitrovich, C. E., Bradshaw, C. P., Poduska, J. M., Hoagwood, K., Buckley, J. A., Olin, S., & Ialongo, N. S. (2008). Maximizing the implementation quality of evidence-based preventive interventions in schools: A conceptual framework. *Advances in School Based Mental Health Promotion, 1*, 6–28.

Dozier, S. G., Schroeder, K., Lee, J., Fulkerson, J. A., & Kubik, M. Y. (2020). The association between parents and children meeting physical activity guidelines. *Journal of pediatric nursing*, *52*, 70–75.

Duck, A. A., Hall, K. C., Klamm, M., Temple, M., & Robinson, J. C. (2020). Physical activity and fitness: The feasibility and preliminary effectiveness of wearable activity tracker technology incorporating altruistic motivation in youth. *Journal for Specialists in Pediatric Nursing*, *26*(1), e12313.

Dudley, D., Okely, A., Pearson, P., & Cotton, W. (2011). A systematic review of the effectiveness of physical education and school sport interventions targeting physical activity, movement skills and enjoyment of physical activity. *European Physical Education Review*, *17*(3), 353–378.

Duncan, M. J., Al-Nakeeb, Y., & Nevill, A. M. (2009). Effects of a 6-week circuit training intervention on body esteem and body mass index in British primary school children. *Body Image*, *6*(3), 216–220.

Duncan, M. J., Eyre, E. L., & Oxford, S. W. (2018). The effects of 10-week integrated neuromuscular training on fundamental movement skills and physical self-efficacy in 6–7-year-old children. *The Journal of Strength & Conditioning Research*, *32*(12), 3348–3356.

Durlak, J. A. (2016). Programme implementation in social and emotional learning: basic issues and research findings. *Cambridge Journal of Education*, *46*(3), 333–345.

Durlak, J. A. & DuPre, E. P. (2008). Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting implementation. *American Journal of Community Psychology, 41*, 327–350.

Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D., & Schellinger, K. B. (2011). The impact of enhancing students' social and emotional learning: A meta□analysis of school□based universal interventions. *Child development, 82*(1), 405–432.

Dyrstad, S. M., Kvalø, S. E., Alstveit, M., & Skage, I. (2018). Physically active academic lessons: Acceptance, barriers and facilitators for implementation. *BMC Public Health, 18*(1), 322.

Eddy, L.H., Wood, M.L., Shire, K.A., Bingham, D.D., Bonnick, E., & Creaser, A. (2019) A systematic review of randomized and case-controlled trials investigating the effectiveness of school-based motor skill interventions in 3- to 12-year-old children. *Child: Care, Health and Development, 45*(6), 773–790.

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *Bmj*, *315*(7109), 629–634.

Elbe, A. M., Wikman, J. M., Zheng, M., Larsen, M. N., Nielsen, G., & Krustrup, P. (2017). The importance of cohesion and enjoyment for the fitness improvement of 8–10-year-old children participating in a team and individual sport school-based physical activity intervention. *European journal of sport science*, *17*(3), 343–350.

Elias, M. J., Wang, M. C., Weissberg, R .P., Zins, J. E., & Walberg, H. J. (2002). The other side of the report card: Student success depends on more than test scores. *American School Boards Journal, 189*(11), 28–31.

Elias, M. J., Zins, J. E., Weissberg, R. P., Frey, K. S., Greenberg, M. T., Haynes, N. M., Kessler, R., Schwab-Stone, M. E., & Shriver, T. P. (1997). *Promoting social and emotional learning: Guidelines for educators*. Alexandria, VA: Association for Supervision and Curriculum Development.

Engel, A. C., Broderick, C. R., van Doorn, N., Hardy, L. L., & Parmenter, B. J. (2018). Exploring the relationship between fundamental motor skill interventions and physical activity levels in children: A systematic review and meta-analysis. *Sports Medicine*, *48*(8), 1845–1857.

Errisuriz, V. L., Golaszewski, N. M., Born, K., & Bartholomew, J. B. (2018). Systematic review of physical education-based physical activity interventions among elementary school children. *The journal of primary prevention*, *39*(3), 303–327.

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

Escartí, A., Gutiérrez, M., Pascual, C., & Llopis, R. (2010). Implementation of the personal and social responsibility model to improve self-efficacy during physical education classes for primary school children. *International Journal of Psychology and Psychological Therapy*, *10*(3), 387–402.

Escartí, A., Llopis-Goig, R., & Wright, P. M. (2018). Assessing the implementation fidelity of a school-based teaching personal and social responsibility program in physical education and other subject areas. *Journal of Teaching in Physical Education, 37*, 12–23.

Evans, R., Murphy, S., & Scourfield, J. (2015). Implementation of a school-based social and emotional learning intervention: understanding diffusion processes within complex systems. *Prevention Science, 16*(5), 754–764.

Figueroa, R., & An, R. (2017). Motor skill competence and physical activity in preschoolers: A review. *Maternal and Child Health Journal, 21*(1), 136–146.

Fisher, A., Reilly, J., Kelly, L. A., Montgomery, C., Williamson, A., Paton, J. Y., & Grant, S. (2005). Fundamental movement skills and habitual physical activity in young children. *Medicine and Science in Sports and Exercise, 37*(4), 684–688.

Flores, M. A. (2020). Preparing teachers to teach in complex settings: Opportunities for professional learning and development. *European Journal of Teacher Education, 43*(3), 297–300.

Fu, Y., & Burns, R. D. (2018). Effect of an active video gaming classroom curriculum on health-related fitness, school day step counts, and motivation in sixth graders. *Journal of Physical Activity and Health*, *15*(9), 644–650.

Gallahue, D., Ozmun, J., & Goodway, J. (2012). *Understanding Motor Development: Infants, Children, Adolescents and Adults*. Seventh Edition. NY: McGraw-Hill.

Goh, T. L., Hannon, J. C., Webster, C., Podlog, L., & Newton, M. (2016). Effects of a take 10! classroom-based physical activity intervention on third-fifth-grade children's on-task behavior. *Journal of Physical Activity & Health*, *13*(7), 712–718.

Goodman, A., Joshi, H., Nasim, B., & Tyler, C., (2015) *Social and emotional skills in childhood and their long term effects on adult life*. UCL, Institute of Education.

Goudas, M., & Magotsiou, E. (2009). The effects of a cooperative physical education program on students' social skills. *Journal of applied sport Psychology*, *21*(3), 356–364.

Greenberg, M. T., Domitrovich, C. E., Weissberg, R. P., & Durlak, J. A. (2017). Social and emotional learning as a public health approach to education. *The future of children*, 13–32.

Greenberg, M. T., Weissberg, R. P., O'Brien, M. U., Zins, J. E., Fredericks, L., Resnik, H., & Elias, M. J. (2003). Enhancing school-based prevention and youth development through coordinated social, emotional, and academic learning. *American Psychologist, 58*(6–7), 466–474.

Grieco, L. A., Jowers, E. M., Errisuriz, V. L., & Bartholomew, J. B. (2016). Physically active vs. sedentary academic lessons: A dose response study for elementary student time on task. *Preventive Medicine, 89*, 98–103.

Griffiths, L. A., & Griffiths, M. A. (2019). Evaluation of a school–community linked physical activity intervention targeting 7-to 12-year-olds: a sociocultural perspective. *American Journal of Health Education*, *50*(2), 112–126.

Grillich, L., Kien, C., Takuya, Y., Weber, M., & Gartlehner, G. (2016). Effectiveness evaluation of a health promotion programme in primary schools: a cluster randomised controlled trial. *BMC public health*, *16*(1), 1–11.

Gustafson, S. L., & Rhodes, R. E. (2006). Parental correlates of physical activity in children and early adolescents. *Sports medicine*, *36*(1), 79–97.

Gu, X., Chen, Y. L., Jackson, A. W., & Zhang, T. (2018). Impact of a pedometer-based goal-setting intervention on children's motivation, motor competence, and physical activity in physical education. *Physical Education and Sport Pedagogy*, *23*(1), 54–65.

Hardman, K. (2011). Global issues in the situation of physical education in schools. In K. Hardman & K. Green (Eds.), *Contemporary issues in physical education* (pp. 11–29). Maidenhead: Meyer & Meyer Sport.

Hardy, L., L. King, P. Espinel, C. Cosgrove, & A. Bauman. (2010). *NSW Schools Physical Activity and Nutrition Survey* (SPANS). Full Report. Sydney: NSW Ministry of Health.

Hassandra, M., & Goudas, M. (2010). An evaluation of a physical education program for the development of students' responsibility. *Hellenic Journal of Psychology*, *7*(3), 275–297.

Hastie, P. A., MacPhail, A., Calderón, A., & Sinelnikov, O. A. (2015). Promoting professional learning through ongoing and interactive support: Three cases within physical education. *Professional development in education*, *41*(3), 452–466.

Hastie, P. A. (2017). Revisiting the national physical education content standards: What do we really know about our achievement of the physically educated/literate person? *Journal of Teaching in Physical Education, 36*(1), 3–19.

Haubenstricker, J., & Seefeldt, V. (1986). *Acquisition of motor skills during childhood. In V. Seefeldt* (Ed.), Physical activity and well-being (pp. 41–92). Reston, VA: American Alliance.

Haywood, K. M., & Getchell, N. (2019). *Life span motor development*. Human kinetics.

Hedges, L. V. (1981). Distribution theory for Glass's estimator of effect size and related estimators. *journal of Educational Statistics*, *6*(2), 107–128.

Heidorn, B., & Centeio, E. (2012). The director of physical activity and staff involvement. *Journal of Physical Education, Recreation & Dance*, *83*(7), 13–26.

Hellison, D. (2011). *Teaching Responsibility through Physical Activity*, 3rd Edn. Champaign, IL: Human Kinetics.

Hennessy, E. A., Johnson, B. T., & Keenan, C. (2019). Best practice guidelines and essential methodological steps to conduct rigorous and systematic meta□reviews. *Applied Psychology: Health and Well-Being, 11*(3), 353–381.

Higgins, J. P. T., & Green, S. (2017). *Cochrane Handbook for Systematic Reviews of Interventions*, Version 5.1. 0. London: The Cochrane Collaboration; 2011.

Higgins, J. P., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M. J., & Welch, V. A. (Eds.). (2019). *Cochrane handbook for systematic reviews of interventions*. John Wiley & Sons.

Hillman, C. H., Pontifex, M. B., Castelli, D. M., Khan, N. A., Raine, L. B., Scudder, M. R., . . . Kamijo, K. (2014). Effects of the FITKids randomized controlled trial on executive control and brain function. *Pediatrics, 134*(4), e1063–1071.

Hintze, J. M., Volpe, R. J., & Shapiro, E. S. (2002). Best practices in the systematic direct observation of student behavior. *Best practices in school psychology, 4*, 993–1006.

Ho, F. K. W., Louie, L. H. T., Wong, W. H. S., Chan, K. L., Tiwari, A., Chow, C. B., ... & Ip, P. (2017). A sports-based youth development program, teen mental health, and physical fitness: An RCT. *Pediatrics*, *140*(4), e20171543.

Howley, D. (2021). Experiences of teaching and learning in K-12 physical education during COVID-19: an international comparative case study. *Physical Education and Sport Pedagogy*, 1–18.

Hsu, C. C., & Sandford, B. A. (2007). The Delphi technique: making sense of consensus. *Practical Assessment, Research, and Evaluation*, *12*(1), 10.

Huang, C., Dannenberg, A. L., Haaland, W., & Mendoza, J. A. (2018). Changes in self-efficacy and outcome expectations from child participation in bicycle Trains for Commuting to and from school. *Health Education & Behavior*, *45*(5), 748–755.

Huedo-Medina, T. B., Sánchez-Meca, J., Marín-Martínez, F., & Botella, J. (2006). Assessing heterogeneity in meta-analysis: Q statistic or I^2 index?. *Psychological methods*, *11*(2), 193.

Hulteen, R. M., Morgan, P. J., Barnett, L. M., Stodden, D. F., & Lubans, D. R. (2018). Development of foundational movement skills: A conceptual model for physical activity across the lifespan. *Sports Medicine, 48*(7), 1533–1540.

International Association for K-12 Online Learning. (2011). *National standards for quality online teaching* (Version 2). Vienna, VA: Author. https://www.inacol.org/resource/inacolnational-standards-for-quality-online-teaching-v2

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

Jennings, P.A. and Frank, J.L. (2015), "Inservice preparation for educators", in Durlak, J., Domotrovich, C.E., Weissberg, R. and Gullotta, T.P. (Eds), *Handbook of Social and Emotional Learning (*pp. 422–438) The Guilford Press, New York, NY.

Jiménez-Díaz, J., Chaves-Castro, K., & Salazar, W. (2019). Effects of different movement programs on motor competence: A systematic review with meta-analysis. *Journal of physical activity and health*, *16*(8), 657–666.

Johnstone, A., Hughes, A. R., Janssen, X., & Reilly, J. J. (2017). Pragmatic evaluation of the Go2Play active play intervention on physical activity and fundamental movement skills in children. *Preventive medicine reports*, *7*, 58–63.

Jones, D., Innerd, A., Giles, E. L., & Azevedo, L. B. (2020). Association between fundamental motor skills and physical activity in the early years: A systematic review and meta-analysis. *Journal of sport and health science*, *9*(6), 542–552.

Jones, S., Bailey, R., Brush, K., & Kahn, J. (2018). *Preparing for effective SEL implementation*. Harvard Graduate School of Education Easel Lab. https://www.wallacefoundation.org/knowledgecenter/Documents/Preparing-for-Effective-SEL-Implementation. pdf.

Jones, S. M., Barnes, S. P., Bailey, R., & Doolittle, E. J. (2017). Promoting social and emotional competencies in elementary school. *The Future of Children, 27*(1), 49–72.

Jones, S. M., & Bouffard, S. M. (2012). Social and emotional learning in schools: From programs to strategies and commentaries. *Social policy report, 26*(4), 1–33.

Jones, T. A., Hawrylak, N., Klintsova, A. Y., & Greenough, W. T. (1998). Brain damage, behavior, rehabilitation, recovery, and brain plasticity. *Mental Retardation and Developmental Disabilities, 4*, 231–237.

Kamijo, K., Pontifex, M. B., O'Leary, K. C., Scudder, M. R., Wu, C. T., Castelli, D. M., & Hillman, C. H. (2011). The effects of an afterschool physical activity program on working memory in preadolescent children. *Developmental Science, 14*(5), 1046–1058.

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*(1), S43–S50.

Killian, C. M., Woods, A. M., Graber, K. C., & Templin, T. J. (2020). Factors associated with high school physical education teachers' adoption of a supplemental online instructional system (iPE). *Journal of Teaching in Physical Education, 40*(1), 136–145.

Kiphard, E. J., & Schilling, F. (2007). *Körperkoordinationstest für kinder: KTK*. Weinheim: Beltz Test GmbH.

Klos, L., Feil, K., Eberhardt, T., & Jekauc, D. (2020). Interventions to Promote Positive Affect and Physical Activity in Children, Adolescents and Young Adults—A Systematic Review. *Sports, 8*(2), 26.

Kontopantelis, E., Springate, D. A., & Reeves, D. (2013). A re-analysis of the Cochrane Library data: the dangers of unobserved heterogeneity in meta-analyses. *PloS one*, *8*(7), e69930.

Kooiman, B. J., & Sheehan, D. P. (2014). The efficacy of exergames played proximally and over the internet on cognitive functioning for online physical education. *American Journal of Distance Education, 28*(4), 280–291.

Kooiman, B. J., Sheehan, D. P., Wesolek, M., & Retegui, E. (2017). Moving online physical education from oxymoron to efficacy. *Sport, Education and Society, 22*(2), 230–246

Kounin, J. S. (1970). *Discipline and group management in classrooms*. New York: Holt, Reinhardt & Winston.

Lai, S. K., Costigan, S. A., Morgan, P. J., Lubans, D. R., Stodden, D. F., Salmon, J., & Barnett, L. M. (2014). Do school-based interventions focusing on physical activity, fitness, or fundamental movement skill competency produce a sustained impact in these outcomes in children and adolescents? A systematic review of follow-up studies. *Sports Medicine, 44*, 67–79.

Lämmle, C., Kobel, S., Wartha, O., Wirt, T., & Steinacker, J. M. (2016). Intervention effects of a school-based health promotion program on children's motor skills. *Journal of Public Health*, *24*(3), 185–192.

Laukkanen, A., Pesola, A. J., Heikkinen, R., Sääkslahti, A. K., & Finni, T. (2015). Family-based cluster randomized controlled trial enhancing physical activity and motor competence in 4–7-year-old children. *PLoS One*, *10*(10), e0141124.

Laurent, C. W. S., Masteller, B., & Sirard, J. (2018). Effect of a suspension-trainer-based movement program on measures of fitness and functional movement in children: A pilot study. *Pediatric exercise science*, *30*(3), 364–375.

Lee, J., Zhang, T., Chu, T. L. A., Gu, X., & Zhu, P. (2020a). Effects of a fundamental motor skill-based afterschool program on children's physical and cognitive health outcomes. *International journal of environmental research and public health*, *17*(3), 733.

Lee, J., Zhang, T., Chu, T. L. A., & Gu, X. (2020b). Effects of a need-supportive motor skill intervention on children's motor skill competence and physical activity. *Children*, *7*(3), 21.

Lendrum, A. & Humphrey, N. (2012). The importance of studying the implementation of interventions in school settings. *Oxford Review of Education, 38*(5), 635–652.

Lewis, K. M., Holloway, S. D., Bavarian, N., Silverthorn, N., DuBois, D. L., Flay, B. R., & Siebert, C. F. (2021). Effects of Positive Action in Elementary School on Student Behavioral and Social-Emotional Outcomes. *The elementary school journal, 121*(4), 635–655.

Lindquist, C. H., Reynolds, K. D., & Goran, M. I. (1999). Sociocultural determinants of physical activity among children. *Preventive Medicine, 29*, 305–312.

Livesey D., Keen J., Rouse J. & White F. (2006). The relationship between measures of executive function, motor performance and externalizing behaviour in 5- and 6-year-old children. *Human Movement Science, 25*(1), 50–64.

Lloyd, M., Saunders, T. J., Bremer, E., & Tremblay, M. S. (2014). Long-term importance of fundamental motor skills: a 20-year follow-up study. *Adapted Physical Activity Quarterly, 31*(1), 67–78.

Logan, S. W., Robinson, L. E., Wilson, A. E., & Lucas, W. A. (2012). Getting the fundamentals of movement: a meta-analysis of the effectiveness of motor skill interventions in children. *Child: Care, Health and Development, 38*(3), 305–315.

Logan, S.W., Webster, E. K., Getchell, N., Pfieffer, K. A., & Robinson, L. E. (2015). Relationship between fundamental motor skill competence and physical activity during childhood and adolescence: A systematic review. *Kinesiology Reviews, 4*(4), 416–426.

Lopes, V. P., Rodrigues, L. P., Maia, J. A. R., & Malina, R. M. (2011). Motor coordination as predictor of physical activity in childhood. *Scandinavian Journal of Medicine and Science in Sports, 21(5)*, 663– 669.

Lopes, V. P., Stodden, D. F., & Rodrigues, L. P. (2017). Effectiveness of physical education to promote motor competence in primary school children. *Physical Education and Sport Pedagogy*, *22*(6), 589–602.

Lorås, H. (2020). The effects of physical education on motor competence in children and adolescents: a systematic review and meta-analysis. *Sports*, *8*(6), 88.

Love, R. Adams, J. & van Sluijs, E.M.F. (2019). Are school-based physical activity interventions effective and equitable? A meta-analysis of cluster randomized controlled trials with accelerometer-assessed activity. *Obesity Review, 20*(6), 859–870.

Lubans, D. R., Morgan, P. J., Cliff, D. P., Barnett, L. M., & Okely, A. D. (2010). Fundamental movement skills in children and adolescents. *Sports Medicine, 40*(12), 1019–1035.

Lwin, M. O., & Malik, S. (2014). Can exergames impart health messages? Game play, framing, and drivers of physical activity among children. *Journal of health communication*, *19*(2), 136–151.

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*(2), 2086–2094.

Martin, R., & Murtagh, E. M. (2015). Preliminary findings of Active Classrooms: An intervention to increase physical activity levels of primary school children during class time. *Teaching and Teacher Education, 52*, 113–127.

Maskell, B., Shapiro, D. R., & Ridley, C. (2004). Effects of Brain Gym on overhand throwing in first grade students: A preliminary investigation. *Physical Educator*, *61*(1), 14–22.

Mayfield, C. A., Child, S., Weaver, R. G., Zarrett, N., Beets, M. W., & Moore, J. B. (2017). Effectiveness of a playground intervention for antisocial, prosocial, and physical activity behaviors. *Journal of school health*, *87*(5), 338–345.

Mayorga-Vega, D., Viciana, J., Cocca, A., & Villén, B. D. R. (2012). Effect of a physical fitness program on physical self-concept and physical fitness elements in primary school students. *Perceptual and Motor Skills*, *115*(3), 984–996.

McClelland, M. M., Tominey, S. L., Schmitt, S. A., & Duncan, R. (2017). SEL interventions in early childhood. *The Future of Children, 27*(1), 33–47.

McDonough, D. J., Liu, W., & Gao, Z. (2020). Effects of Physical Activity on Children's Motor Skill Development: A Systematic Review of Randomized Controlled Trials. *BioMed research international*, *2020*.

McGuinness, L. A., & Higgins, J. P. (2021). Risk□of□bias VISualization (robvis): An R package and Shiny web app for visualizing risk□of□bias assessments. *Research Synthesis Methods*, *12*(1), 55–61.

McKenzie, T. L., Alcaraz, J. E., Sallis, J. F., & Faucette, F. N. (1998). Effects of a physical education program on children's manipulative skills. *Journal of Teaching in Physical education*, *17*(3), 327–341.

McKenzie, T. L., & Smith, N. J. (2017). Studies of physical education in the United States using SOFIT: A review. *Research Quarterly for Exercise and Sport, 88*(4), 492–502.

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*(1), 13–29.

McWhannell, N., Foweather, L., Graves, L. E., Henaghan, J. L., Ridgers, N. D., & Stratton, G. (2018). From surveillance to intervention: overview and baseline findings for the active city of liverpool active schools and SportsLinx (A-class) project. *International journal of environmental research and public health*, *15*(4), 582.

Mendelson, T., Greenberg, M. T., Dariotis, J. K., Gould, L. F., Rhoades, B. L., & Leaf, P. J. (2010). Feasibility and preliminary outcomes of a school-based mindfulness intervention for urban youth. *Journal of abnormal child psychology*, *38*(7), 985–994.

Mercier, K., Centeio, E., Garn, A., Erwin, H., Marttinen, R., & Foley, J. (2021). Physical education teachers' experiences with remote instruction during the initial phase of the COVID-19 pandemic. *Journal of Teaching in Physical Education, 40*(2), 337–342.

Meyers, D. C., Durlak, J. A., & Wandersman, A. (2012). The quality implementation framework: A synthesis of critical steps in the implementation process. *American Journal of Community Psychology, 50*, 462–480.

Michael' R. D. Webster, C. A., Egan, C. A., Nilges, L., Brian, A., Johnson, R., & Carson, R. L. (2019). Facilitators and barriers to movement integration in elementary classrooms: A systematic review. *Research Quarterly for Exercise and Sport*, *90*(2), 151–162.

Miller, A., Christensen, E., Eather, N., Gray, S., Sproule, J., Keay, J., & Lubans, D. (2016). Can physical education and physical activity outcomes be developed simultaneously using a game-centered approach? *European Physical Education Review*, *22*(1), 113–133.

Miller, A., Christensen, E. M., Eather, N., Sproule, J., Annis-Brown, L., & Lubans, D. R. (2015). The PLUNGE randomized controlled trial: Evaluation of a games-based physical activity professional learning program in primary school physical education. *Preventive medicine*, *74*, 1–8.

Miyake, A. & Friedman, N. P. (2012). The nature and organization of individual differences in executive functions: four general conclusions. *Current Directions in Psychological Science, 21*(1), 8–14.

Mohnsen, B. (2012). Implementing online physical education. *Journal of Physical Education, Recreation & Dance, 83*(2), 42–47.

Moon, J., & Webster, C. A. (2019). MI (my) wheelhouse: A movement integration progression framework for elementary classroom teachers. *Journal of Physical Education, Recreation and Dance, 90*(7), 38e45.

Moon, J., Webster, C. A., Herring, J., & Egan, C. A. (2020). Relationships between systematically observed movement integration and classroom management in elementary schools. *Journal of Positive Behavior Interventions*.

Murcia, J. A. M., & Latorre, F. S. (2016). The effects of autonomy support in physical education classes. *RICYDE. Revista Internacional de Ciencias del Deporte*, *12*(43), 79–89.

Morgan, P. J., Barnett, L. M., Cliff, D. P., Okely, A. D., Scott, H. A., Cohen, K. E., & Lubans, D. R. (2013). Fundamental movement skill interventions in youth: A systematic review and meta-analysis. *Pediatrics, 132*, e1361–e1383.

Mosier, B., & Lynn, S. (2012). An initial exploration of a virtual personal fitness course. Online *Journal of Distance Learning Administration, 15*(3). https://www.westga.edu/~distance/ojdla/fall153/ mosier_lynn153.html

Mullender-Wijnsma, M. J., Hartman, E., de Greeff, J. W., Doolaard, S., Bosker, R. J., & Visscher, C. (2016). Physically active math and language lessons improve academic achievement: A cluster randomized controlled trial. *Pediatrics, 137*(3), e20152743.

Mulvey, K. L., Taunton, S., Pennell, A., & Brian, A. (2018). Head, toes, knees, skip! Improving preschool children's executive function through a motor competence intervention. *Journal of Sports and Exercise Psychology, 40*(5), 233–239.

Murano, D., Sawyer, J. E., & Lipnevich, A. A. (2020). A meta-analytic review of preschool social and emotional learning interventions. *Review of Educational Research, 90*(2), 227–263.

Myer, G. D., Faigenbaum, A. D., Edwards, N. M., Clark, J. F., Best, T. M., & Sallis, R. E. (2015). Sixty minutes of what? A developing brain perspective for activating children with an integrative exercise approach. *British Journal of Sports Medicine, 49*(23), 1510–1516.

Nathan, N., Sutherland, R., Beauchamp, M. R., Cohen, K., Hulteen, R. M., Babic, M., Wolfenden, L., & Lubans, D. R. (2017). Feasibility and efficacy of the Great Leaders Active StudentS (GLASS) program on children’s physical activity and object control skill competency: A non-randomised trial. *Journal of science and medicine in sport*, *20*(12), 1081–1086.

National Association for Sport and Physical Education. (2008). Comprehensive school physical activity programs [position statement]. https://eric.ed.gov/?id=ED541610

National Association for Sport and Physical Education. (2007). Initial guidelines for online physical education [Position paper]. http:// www.dpi.state.nc.us/docs/curriculum/healthfulliving/resources/onlinepeguidelines .pdf

National Physical Activity Plan Alliance. (2018). *The 2018 United States Report Card on Physical Activity for Children and Youth*. National Physical Activity Plan Alliance.

Nesbitt, D., Fisher, J., & Stodden, D. F. (2021). Appropriate instructional practice in physical education: A systematic review of literature from 2000 to 2020. *Research Quarterly for Exercise and Sport*, *92*(2), 235–247.

Newell, K. M. (2020). What are Fundamental Motor Skills and What is Fundamental About Them? *Journal of Motor Learning and Development*, *8*(2), 280–314.

Norman, D. A., & Shallice, T. (1986). Attention to action: Willed and automatic control of behavior. In R. J. Davidson, G. E. Schwartz, & D. Shapiro (Eds.), *Consciousness and self-regulation* (Vol. 4, pp. 1-18). New York: Plenum Press.

Norris, E., Shelton, N., Dunsmuir, S., Duke-Williams, O., & Stamatakis, E. (2015). Physically active lessons as physical activity and educational interventions: A systematic review of methods and results. *Preventive Medicine, 72*, 116e125.

Okely, A. D., Hardy, L. L., Batterham, M., Pearson, P., McKeen, K., & Puglisi, L. (2017). Promoting motor skills in low-income, ethnic children: The Physical Activity in Linguistically Diverse Communities (PALDC) nonrandomized trial. *Journal of science and medicine in sport*, *20*(11), 1008–1014.

Olive, C., McCullick, B. A., Tomporowski, P., Gaudreault, K. L., & Simonton, K. (2020). Effects of an After-School Program Focused on Physical Activity and Social–Emotional Learning. *Journal of Youth Development*, *15*(6), 292–305.

Olive, L. S., Byrne, D., Cunningham, R. B., Telford, R. M., & Telford, R. D. (2019). Can physical education improve the mental health of children? The LOOK study cluster-randomized controlled trial. *Journal of educational psychology*, *111*(7), 1331.

Orendorff, K., Webster, C. A., Mîndrilă, D., Cunningham, K. M., Doutis, P., Dauenhauer, B., & Stodden, D. F. (2020). Principals' involvement in comprehensive school physical activity programmes: A social-ecological perspective. *European Physical Education Review*.

Osher, D., Kidron, Y., Brackett, M., Dymnicki, A., Jones, S., & Weissberg, R. P. (2016). Advancing the science and practice of social and emotional learning: Looking back and moving forward. *Review of Research in Education*, *40*(1), 644–681.

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., ... & Moher, D. (2021). Updating guidance for reporting systematic reviews: development of the PRISMA 2020 statement. *Journal of Clinical Epidemiology*, *134*, 103–112.

Palmer, K. K., Chinn, K. M., & Robinson, L. E. (2017). Using achievement goal theory in motor skill instruction: a systematic review. *Sports Medicine*, *47*(12), 2569–2583.

Payton, J. W., Wardlaw, D. M., Graczyk, P. A., Bloodworth, M. R., Tompsett, C. J., & Weissberg, R. P. (2000). Social and emotional learning: A framework for promoting mental health and reducing risk behavior in children and youth. *Journal of School Health, 70*(5), 179–185.

Pérez-Ordás, R., Pozo, P., & Grao-Cruces, A. (2020). Effects on aggression and social responsibility by teaching personal and social responsibility during physical education. *Journal of Physical Education and Sport*, *20*(4), 1832–1838.

Pesce, C., & Ben-Soussan, T. D. (2016). ""Cogito ergo sum" or "ambulo ergo sum"? New perspectives in developmental exercise and cognition research," in *Exercise-Cognition Interaction. Neuroscience Perspectives*, ed T. McMorris (London: Elsevier), 251–282.

Pesce, C., Croce, R., Ben-Soussan, T. D., Vazou, S., McCullick, B., Tomporowski, P. D., & Horvat, M. (2019). Variability of practice as an interface between motor and cognitive development. *International Journal of Sport and Exercise Psychology, 17*(2), 133–152.

Pesce, C., Crova, C., Marchetti, M., Struzzolino, I., Masci, I., Vannozzi, G., & Forte, R. (2013). Searching for cognitively optimal challenge point in physical activity for children with typical and atypical motor development. *Mental Health and Physical Activity, 6*(3), 172–180.

Pesce, C., Faigenbaum, A., Crova, C., Marchetti, R., & Bellucci, M. (2013). Benefits of multi-sports physical education in the elementary school context. *Health Education Journal*, *72*(3), 326–336.

Pesce, C., Leone, L., Motta, A., Marchetti, R., & Tomporowski, P. D. (2016). From efficacy to effectiveness of a 'whole' child initiative of physical activity promotion. *Translational Journal of the American College of Sports Medicine, 1*(3), 18–29.

Pesce, C., Vazou, S., Benzing, V., Álvarez-Bueno, C., Anzeneder, S., Mavilidi, M. F., Leone, L., & Schmidt, M. (2021). Effects of chronic physical activity on cognition across the lifespan: a systematic meta-review of randomized controlled trials and realist synthesis of contextualized mechanisms, *International Review of Sport and Exercise Psychology*.

Platvoet, S. W., Elferink-Gemser, M. T., Kannekens, R., de Niet, M., & Visscher, C. (2016). Four weeks of goal-directed learning in primary physical education classes. *Perceptual and motor skills*, *122*(3), 871–885.

Pozo, P., Grao-Cruces, A., & Pérez-Ordás, R. (2018). Teaching personal and social responsibility model-based programmes in physical education: A systematic review. *European Physical Education Review*, *24*(1), 56–75.

Prevention CfDCa. Staff Involvement 2020 [Available from: https://www.cdc. gov/healthyschools/physicalactivity/staff-involvement.htm.

Pulling Kuhn, A., Stoepker, P., Dauenhauer, B., & Carson, R. L. (2021). A systematic review of multi-component comprehensive school physical activity program (CSPAP) interventions. *American Journal of Health Promotion*.

Quarmby, T., Daly-Smith, A., & Kime, N. (2019). 'You get some very archaic ideas of what teaching is…': primary school teachers' perceptions of the barriers to physically active lessons. *Education 3-13, 47*(3), 1–14.

Quested, E., Ntoumanis, N., Thøgersen-Ntoumani, C., Hagger, M. S., & Hancox, J. E. (2017). Evaluating quality of implementation in physical activity interventions based on theories of motivation: current challenges and future directions. *International Review of Sport and Exercise Psychology*, *10*(1), 252–269.

Ransdell, L. B., Rice, K., Snelson, C., & Decola, J. (2008). Online health-related fitness courses. *The Journal of Physical Education, Recreation & Dance, 79*(1), 45–52.

Rasberry, C. N., Lee, S. M., Robin, L., Laris, B. A., Russell, L. A., Coyle, K. K., & Nihiser, A. J. (2011). The association between school-based physical activity, including physical education, and academic performance: a systematic review of the literature. *Preventive medicine, 52*, S10-S20.

Ré, A. H., Logan, S. W., Cattuzzo, M. T., Henrique, R. S., Tudela, M. C., & Stodden, D. F. (2018). Comparison of motor competence levels on two assessments across childhood. *Journal of sports sciences*, *36*(1), 1–6.

Resaland G. K., Moe V. F., Bartholomew J. B., Andersen L. B., McKay H. A., Anderssen S. A., & Aadland, E. (2018). Gender-specific effects of physical activity on children's academic performance: the active smarter kids cluster randomized controlled trial. *Preventive Medicine, 106*, 171–176.

Rhea, D. J. (2011). Virtual physical education in the K-12 setting. *Journal of Physical Education, Recreation & Dance, 82*(1), 5–50.

Riethmuller, A. M., Jones, R. A., & Okely, A. D. (2009). Efficacy of interventions to improve motor development in young children: A systematic review. *Pediatrics, 124*, e782–e792.

Riggs, N. R., Jahromi, L. B., Razza, R. P., Dillworth-Bart, J. E., & Mueller, U. (2006). Executive function and the promotion of social-emotional competence. *Journal of Applied Developmental Psychology, 27*(4), 300–309.

Riley, N., Lubans, D. R., Morgan, P. J., & Young, M. (2015). Outcomes and process evaluation of a programme integrating physical activity into the primary school mathematics curriculum: The EASY Minds pilot randomized controlled trial. *Journal of Science and Medicine in Sport, 18*(6), 656–661.

Rink, J. E. (2020). *Teaching physical education for learning* (8th ed.). New York, NY: McGraw-Hill.

Robinson, L. E. & Goodway, J. D. (2009). Instructional climates in preschool children who are at-risk. Part I: object-control skill development. *Research Quarterly for Exercise and Sport, 80*(3), 533–542.

Robinson, L. E., Palmer, K. K., & Meehan, S. K. (2017). Dose–response relationship: the effect of motor skill intervention duration on motor performance. *Journal of Motor Learning and Development*, *5*(2), 280–290.

Robinson, L. E., Stodden, D. F., Barnett, L. M., Lopes, V. P., Logan, S. W., Rodrigues, L. P., & D'Hondt, E. (2015). Motor competence and its effect on positive developmental trajectories of health. *Sports Medicine*, *45*(9), 1273–1284.

Roebers C. M. & Kauer M. (2009). Motor and cognitive control in a normative sample of 7-year-olds. *Developmental Science, 12*(1), 175–181.

Rudd, J. R., Barnett, L. M., Farrow, D., Berry, J., Borkoles, E., & Polman, R. (2016). Effectiveness of a 16 week gymnastics curriculum at developing movement competence in children. *Journal of science and medicine in sport*, *20*(2), 164–169.

Rudd, J. R., Barnett, L. M., Farrow, D., Berry, J., Borkoles, E., & Polman, R. (2017). The impact of gymnastics on children's physical self-concept and movement skill development in primary schools. *Measurement in physical education and exercise science*, *21*(2), 92–100.

Russ, L. B., Webster, C. A., Beets, M. W., Egan, C., Weaver, R. G., Harvey, R., & Phillips, D. S. (2017). Development of the system for observing student movement in academic routines and transitions (SOSMART). *Health Education & Behavior*, *44*(2), 304–315.

Russ, L. B., Webster, C. A., Beets, M. W., & Phillips, D. S. (2015). Systematic review and meta- analysis of multi-component interventions through schools to increase physical activity. *Journal of Physical Activity & Health, 12*(10), 1436–1446.

Schmidt, M., Benzing, V., & Kamer, M. (2016). Classroom-based physical activity breaks and children's attention: cognitive engagement works! *Frontiers in psychology, 7*, 1474.

Serrien, D. J., Ivry, R. B., & Swinnen, S. P. (2007). The missing link between action and cognition. *Progress in Neurobiology, 82*, 95–107.

Shachar, K., Ronen-Rosenbaum, T., Rosenbaum, M., Orkibi, H., & Hamama, L. (2016). Reducing child aggression through sports intervention: The role of self-control skills and emotions. *Children and youth services review*, *71*, 241–249.

SHAPE America. (2018). *Guidelines for K-12 online physical education*. [Guidance document]. Reston, VA: Author. https://www.shapeamerica.org/uploads/pdfs/2020/guidelines/Online-PE-Guidance-Document.pdf

SHAPE America-Society of Health and Physical Educators. (2014). *National standards & grade-level outcomes for K-12 physical education*. Human Kinetics, Incorporated.

Shen, B., Centeio, E., Garn, A., Martin, J., Kulik, N., Somers, C., & McCaughtry, N. (2018). Parental social support, perceived competence and enjoyment in school physical activity. *Journal of sport and health science, 7*(3), 346–352.

Silveira, D. S., Lemos, L. F. G. B. F., Tassitano, R. M., Cattuzzo, M. T., Feitoza, A. H. P., Aires, L. M. S. M. C., ... & Martins, C. M. D. L. (2018). Effect of a pilot multi-component intervention on motor performance and metabolic risks in overweight/obese youth. *Journal of sports sciences*, *36*(20), 2317–2326.

Simonton, K. L., & Shiver, V. N. (2021). Examination of elementary students' emotions and personal and social responsibility in physical education. *European Physical Education Review*.

Skowroński, W., Skowrońska, M., Rutkowska, I., Bednarczuk, G., Kaźmierska-Kowalewska, K. M., & Marszałek, J. (2019). The effects of extracurricular physical education classes on gross motor development in primary school children–pilot study. *Biomedical Human Kinetics*, *11*(1), 136–143.

Society of Health and Physical Educators (SHAPE) America. (2015). *Comprehensive school physical activity programs: Helping all students achieve 60 minutes of physical activity each day*. https://www.shapeamerica.org//advocacy/positionstatements/pa/upload/Comprehensive-School-Physical-Activity-Programs-Helping-All-Students-Log-60-Minutes-of-Physical-Activity-Eeach-Day.pdf

Sterne, J. A., Hernán, M. A., Reeves, B. C., Savović, J., Berkman, N. D., Viswanathan, M., . . . Boutron, I. (2016). ROBINS-I: A tool for assessing risk of bias in non-randomised studies of interventions. *BMJ, 355*, i4919.

Sterne, J. A., Savovic, J., Page, M. J., Elbers, R. G., Blencowe, N. S., Boutron, I., . . . Higgins, J. P. T. (2019). RoB 2: A revised tool for assessing risk of bias in randomised trials. *BMJ, 366*, l4898.

Stöckel, T., & Hughes, C. M. (2016). The relation between measures of cognitive and motor functioning in 5- to 6-year-old children. *Psychological Research, 80*, 543–554.

Stodden, D. F., Goodway, J. D., Langendorfer, S. J., Roberton, M. A., Rudisill, M. E., Garcia, C., & Garcia, L. E. (2008). A developmental perspective on the role of motor skill competence in physical activity: An emergent relationship. *Quest, 60*(2), 290–306.

Stodden, D. F., Howard, R., Faigenbaum, A. D., Richardson, C., Meadors, L., Moore, W., & Cullen-Carroll, P. (2012). Promoting integrative youth physical development in the United States. *Professional Strength and Conditioning Journal, 26*, 10–18.

Ströhle, A. (2009). Physical activity, exercise, depression and anxiety disorders. *Journal of Neural Transmission, 116*(6), 777–784.

Strong, W. B., Malina, R. M., Blimkie, C. J., Daniels, S. R., Dishman, R. K., Gutin, B., ... & Rowland, T. (2005). Evidence based physical activity for school-age youth. *The Journal of Pediatrics, 146*(6), 732–737.

Tannehill, D., Demirhan, G., Čaplová, P., & Avsar, Z. (2021). Continuing professional development for physical education teachers in Europe. *European Physical Education Review*, *27*(1), 150–167.

Tester, G., Ackland, T. R., & Houghton, L. (2014). A 30-year journey of monitoring fitness and skill outcomes in physical education: Lessons learned and a focus on the future. *Advances in Physical Education, 4*(3), 127–137.

Trost, S.G., Pate, R.R., Sallis, J.F., Freedson, P., Taylor, W.C., Dowda, M., & Sirard, J.R. (2002). Age and gender differences in objectively measured physical activity in youth. *Medicine and Science in Sports and Exercise, 34*(2), 350–355.

Turoff, M., & Linstone, H. A. (2002). The Delphi method-techniques and applications.

Ulrich, D.A. (2000). *Test of Gross Motor Development* (2nd ed.). Pro-ED.

Ulrich, D. A. (2019). *Test of gross motor development* (3rd ed.). Pro-ED.

Utesch, T., Bardid, F., Büsch, D., & Strauss, B. (2019). The relationship between motor competence and physical fitness from early childhood to early adulthood: A meta-analysis. *Sports Medicine*, *49*(4), 541–551.

Van Capelle, A., Broderick, C. R., Van Doorn, N., & Parmenter, B. J. (2017). Interventions to improve fundamental motor skills in pre-school aged children: a systematic review and meta-analysis. *Journal of Science and Medicine in Sport, 20*(7), 658–666.

Van den Berg, V., Salimi, R., De Groot, R. H., Jolles, J., Chinapaw, M. J., & Singh, A. S. (2017). "It'sa battle… you want to do it, but how will you get it done?": Teachers' and principals' perceptions of implementing additional physical activity in school for academic performance. *International journal of environmental research and public health*, *14*(10), 1160.

Van der Fels, I. M., TeWierike, S. C., Hartman, E., Elferink-Gemser, M. T., Smith, J., & Visscher, C. (2015). The relationship between motor skills and cognitive skills in 4–16 year old typically developing children: A systematic review. *Journal of Science and Medicine in Sport*, *18*(6), 697–703.

Van der Mars, H. (1989). Observer reliability: Issues and procedures. In P. W. Darst. D. Zakrajsek, & V. H. Mancini. *Analyzing Physical Education and Sport Instruction* (Eds.), (2nd edn.) (pp. 53-80). Champaign, IL: Human Kinetics.

Van der Niet, A.G., Smith, J., Oosterlaan, J., Scherder, E. J. A., Hartman, E., & Visscher, C. (2016). Effects of a cognitively demanding aerobic intervention during recess on children's physical fitness and executive functioning. *Pediatric Exercise Science, 28*(1), 64–70.

Vazou, S., Pesce, C., Lakes, K., & Smiley-Oyen, A. (2019). More than one road leads to Rome: A narrative review and meta-analysis of physical activity intervention effects on cognition in youth. *International Journal of Sport and Exercise Psychology, 17*(2), 153–178.

Veldman, S. L., Jones, R. A., & Okely, A. D. (2016). Efficacy of gross motor skill interventions in young children: An updated systematic review. *BMJ Open Sport & Exercise Medicine, 2*(1), e000067.

Ward, P. (2020). Core practices for teaching physical education: Recommendations for teacher education. *Journal of Teaching in Physical Education*, *40*(1), 98–108.

Watson, A., Timperio, A., Brown, H., Best, K., & Hesketh, K. D. (2017). Effect of classroombased physical activity interventions on academic and physical activity outcomes: a systematic review and meta-analysis. *International Journal of Behavioral Nutrition and Physical Activity, 14*(1), 114.

Watson, J.F., Murin, A., Vashaw, L., Gemin, G., & Rapp, C. (2012). Keeping pace with K-12 online learning: An annual review of policy and practice. http://www.kpk12.com

Webster, C. A., Glascoe, G., Moore, C., Dauenhauer, B., Egan, C. A., Russ, L. B., … & Buschmeier, C. (2020). Recommendations for administrators' involvement in school-based health promotion: A scoping review. *International Journal of Environmental Research and Public Health, 17*(17), 6249.

Webster, C. A., Mîndrilă, D., & Weaver, G. (2011). The influence of state motivation, content relevance and affective learning on high school students' intentions to use class content following completion of compulsory physical education*. Journal of Teaching in Physical Education, 30*(3), 231–247.

Webster, C. A., Rink, J. E., Carson, R. L., Moon, J., & Gaudreault, K. L. (2020). The comprehensive school physical activity program model: A proposed illustrative supplement to help move the needle on youth physical activity. *Kinesiology Review*, *9*(2), 112–121.

Webster, C. A., Russ, L., Vazou, S., Goh, T., & Erwin, H. (2015). Integrating movement in academic classrooms: Understanding, applying, and advancing the knowledge base. *Obesity Reviews, 16*(8), 691–701.

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

Wick, K., Leeger-Aschmann, C.S., Monn, N.D., Radtke, T., Ott, L.V., Rebholz, C.E., . . . Kriemler, S. (2017). Interventions to promote fundamental movement skills in childcare and kindergarten: A systematic review and meta-analysis. *Sports Medicine, 47*(10), 2045–2068.

Wigelsworth, M., Lendrum, A., Oldfield, J., Scott, A., Ten-Bokkel, I., Tate, K., & Emery, C. (2016). The impact of trial stage, developer involvement and international transferability on universal social and emotional learning programme outcomes*: a meta-analysis. Cambridge Journal of Education, 46(34),* 345–376.

Wigelsworth, M., Verity, L., Mason, C., Humphrey, N., Qualter, P., & Troncoso, P. (2020). Programmes to Practices Identifying Effective, Evidence-Based Social and Emotional Learning Strategies for Teachers and Schools: Evidence Review. *Education Endowment Foundation.*

Williams, H. G., Pfeiffer, K. A., O'neill, J. R., Dowda, M., McIver, K. L., Brown, W. H., & Pate, R. R. (2008). Motor skill performance and physical activity in preschool children. *Obesity*, *16*(6), 1421–1426.

Williams, L. M. (2013). *A case study of virtual physical education teachers' experiences and perspectives of online teaching* [Unpublished doctoral dissertation]. University of South Florida.

Williams, L., Martinasek, M., Karone, K., & Sanders, S. (2020). High school students' perceptions of traditional and online health and physical education courses. *Journal of School Health, 90*(3), 234–244.

Wong, W. W., Ortiz, C. L., Stuff, J. E., Mikhail, C., Lathan, D., Moore, L. A., ... & Smith, E. O. B. (2016). A community-based healthy living promotion program improved self-esteem among minority children. *Journal of pediatric gastroenterology and nutrition*, *63*(1), 106–112.

Wood, P. (2020). Emotional Intelligence and Social and Emotional Learning:(Mis) Interpretation of Theory and Its Influence on Practice. *Journal of Research in Childhood Education, 34*(1), 153–166.

World Health Organization. (2019). *Global action plan on physical activity 2018-2030: more active people for a healthier world*. World Health Organization.

Worrell, A., Hushman, G., L. Gaudreault, K., Mallett, L., & Hushman, C. (2020). Merging Social and Emotional Learning with Comprehensive School Physical Activity Programming in an Elementary School. *Journal of Physical Education, Recreation & Dance, 91*(6), 36–40.

Wrotniak, B. H., Epstein, L. H., Dorn, J. M., Jones, K. E., & Kondilis, V. A. (2006). The relationship between motor proficiency and physical activity in children. *Pediatrics, 118*(6), 1758–1765.

Zysset, A. E., Kakebeeke, T. H., Messerli-Bürgy, N., Meyer, A. H., Stülb, K., Leeger-Aschmann, C. S., ... & Jenni, O. G. (2018). The validity of parental reports on motor skills performance level in preschool children: a comparison with a standardized motor test. *European journal of pediatrics*, *177*(5), 715–722.

APPENDIX A

THE PRISMA 2020 MAIN CHECKLIST FOR STUDY 1

Table A.1 The PRISMA 2020 Checklist for Study 1

Section and Topic	Item #	Checklist item	Location where item is reported
TITLE			
Title	1	Identify the report as a systematic review.	8
ABSTRACT			
Abstract	2	See the PRISMA 2020 for Abstracts checklist.	N/A
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of existing knowledge.	10-11
Objectives	4	Provide an explicit statement of the objective(s) or question(s) the review addresses.	10
METHODS			
Eligibility criteria	5	Specify the inclusion and exclusion criteria for the review and how studies were grouped for the syntheses.	10-11
Information sources	6	Specify all databases, registers, websites, organisations, reference lists and other sources searched or consulted to identify studies. Specify the date when each source was last searched or consulted.	169
Search strategy	7	Present the full search strategies for all databases, registers and websites, including any filters and limits used.	12-13
Selection process	8	Specify the methods used to decide whether a study met the inclusion criteria of the review, including how many reviewers screened each record and each report retrieved, whether they worked independently, and if applicable, details of automation tools used in the process.	N/A
Data collection process	9	Specify the methods used to collect data from reports, including how many reviewers collected data from each report, whether they worked independently, any processes for obtaining or confirming data from study	12-13

Section and Topic	Item #	Checklist item	Location where item is reported
		investigators, and if applicable, details of automation tools used in the process.	
Data items	10a	List and define all outcomes for which data were sought. Specify whether all results that were compatible with each outcome domain in each study were sought (e.g. for all measures, time points, analyses), and if not, the methods used to decide which results to collect.	N/A
	10b	List and define all other variables for which data were sought (e.g. participant and intervention characteristics, funding sources). Describe any assumptions made about any missing or unclear information.	12-13
Study risk of bias assessment	11	Specify the methods used to assess risk of bias in the included studies, including details of the tool(s) used, how many reviewers assessed each study and whether they worked independently, and if applicable, details of automation tools used in the process.	15
Effect measures	12	Specify for each outcome the effect measure(s) (e.g. risk ratio, mean difference) used in the synthesis or presentation of results.	13-14
Synthesis methods	13a	Describe the processes used to decide which studies were eligible for each synthesis (e.g. tabulating the study intervention characteristics and comparing against the planned groups for each synthesis (item #5)).	12-13
	13b	Describe any methods required to prepare the data for presentation or synthesis, such as handling of missing summary statistics, or data conversions.	13-14
	13c	Describe any methods used to tabulate or visually display results of individual studies and syntheses.	174-184
	13d	Describe any methods used to synthesize results and provide a rationale for the choice(s). If meta-analysis was performed, describe the model(s), method(s) to identify the presence and extent of statistical heterogeneity, and software package(s) used.	13-14
	13e	Describe any methods used to explore possible causes of heterogeneity among study results (e.g. subgroup analysis, meta-regression).	13-14
	13f	Describe any sensitivity analyses conducted to assess robustness of the synthesized results.	N/A

Section and Topic	Item #	Checklist item	Location where item is reported
Reporting bias assessment	14	Describe any methods used to assess risk of bias due to missing results in a synthesis (arising from reporting biases).	15
Certainty assessment	15	Describe any methods used to assess certainty (or confidence) in the body of evidence for an outcome.	N/A
RESULTS			
Study selection	16a	Describe the results of the search and selection process, from the number of records identified in the search to the number of studies included in the review, ideally using a flow diagram.	16
	16b	Cite studies that might appear to meet the inclusion criteria, but which were excluded, and explain why they were excluded.	N/A
Study characteristics	17	Cite each included study and present its characteristics.	174-184
Risk of bias in studies	18	Present assessments of risk of bias for each included study.	16
Results of individual studies	19	For all outcomes, present, for each study: (a) summary statistics for each group (where appropriate) and (b) an effect estimate and its precision (e.g. confidence/credible interval), ideally using structured tables or plots.	16-25
Results of syntheses	20a	For each synthesis, briefly summarise the characteristics and risk of bias among contributing studies.	25-28
	20b	Present results of all statistical syntheses conducted. If meta-analysis was done, present for each the summary estimate and its precision (e.g. confidence/credible interval) and measures of statistical heterogeneity. If comparing groups, describe the direction of the effect.	25-28
	20c	Present results of all investigations of possible causes of heterogeneity among study results.	17-25
	20d	Present results of all sensitivity analyses conducted to assess the robustness of the synthesized results.	16-28
Reporting biases	21	Present assessments of risk of bias due to missing results (arising from reporting biases) for each synthesis assessed.	40-41
Certainty of evidence	22	Present assessments of certainty (or confidence) in the body of evidence for each outcome assessed.	N/A
DISCUSSION			

Section and Topic	Item #	Checklist item	Location where item is reported
Discussion	23a	Provide a general interpretation of the results in the context of other evidence.	28-36
	23b	Discuss any limitations of the evidence included in the review.	32-33
	23c	Discuss any limitations of the review processes used.	36-37
	23d	Discuss implications of the results for practice, policy, and future research.	37-38
OTHER INFORMATION			
Registration and protocol	24a	Provide registration information for the review, including register name and registration number, or state that the review was not registered.	10
	24b	Indicate where the review protocol can be accessed, or state that a protocol was not prepared.	N/A
	24c	Describe and explain any amendments to information provided at registration or in the protocol.	10
Support	25	Describe sources of financial or non-financial support for the review, and the role of the funders or sponsors in the review.	N/A
Competing interests	26	Declare any competing interests of review authors.	N/A
Availability of data, code and other materials	27	Report which of the following are publicly available and where they can be found: template data collection forms; data extracted from included studies; data used for all analyses; analytic code; any other materials used in the review.	169

APPENDIX B

ARTICLE SEARCH STRATEGY AND KEY TERMS FOR

STUDY 1 MOTOR COMPETENCE OUTCOMES

Table B.1 Search Strategy and Key Terms for Study 1

Population	Study Design	Intervention Focus	Motor skill competence related outcomes	Database
Child*[tiab] OR Child*[mh] OR OR Boys[tiab] OR Girls[tiab] OR Student[tiab] OR Student*[tiab]	Intervention[tiab] OR Experiment[tiab] OR Program*[tiab] OR Evaluation[tiab] OR Trial[tiab] OR Random*[tiab] OR Clinic*[tiab] OR “Controlled trial”[tiab]	“Physical activity” [tiab] OR Exercise[tiab]	“Motor competence” [tiab] OR “Motor skills”[tiab] OR “Movement skill”[tiab] OR Coordination[tiab] OR “Motor development” [tiab] OR “Motor skill competence”[tiab] OR “Locomotor skills” [tiab] OR “Fundamental motor skills” [tiab] OR “Motor proficiency”[tiab] OR “Object control”[tiab]	PubMed/ Medline
Youth* or Child* or Student* or Boys or Girls or Teenager* or Teen*	Intervention or Experiment or Program or Evaluation or Trial or Random* or Clinic* or Controlled trial	Physical activity or Exercise	Motor competence or Motor skills or Movement skill or Coordination or Motor development or Motor skill competence or Locomotor skills or Fundamental motor skills or Motor proficiency or Object control	PsycINFO SPORTDiscus ERIC CINAHL Web of Science

(Youth*:ab,ti OR Child*:ab,ti OR Student*:ab,ti OR teenager*:ab,ti OR teen*:ab,ti OR Boys:ab,ti OR Girls:ab,ti) AND (Intervention:ab,ti OR Experiment:ab,ti OR Program*:ab,ti OR Evaluation:ab,ti OR Trial:ab,ti OR Random*:ab,ti OR Clinic*:ab,ti OR 'Controlled trial':ab,ti) AND ('Physical activity':ab,ti OR Exercise:ab,ti) AND ('Motor competence':ab,ti OR 'Motor skills':ab,ti OR 'Movement skill':ab,ti OR Coordination:ab,ti OR 'Motor development':ab,ti OR 'Motor skill competence':ab,ti OR 'Locomotor skills':ab,ti OR 'Fundamental motor skills':ab,ti OR 'Motor proficiency':ab,ti OR 'Object control':ab,ti)	Embase

APPENDIX C

OVERVIEW OF CHARACTERISTICS OF INCLUDED STUDIES

EXAMINING MOTOR COMPETENCE

Table C.1 Characteristics of Included Studies for Study 1

Study	Study Design	CSPAP Component	Population, Sample size, Sex, and Age	Duration & Dose	Description of Intervention	Intervention Deliverer	Measurement of MC	Fidelity	Main Findings in MC Outcomes as Reported
Bolger et al. (2019).	QE	PE	*N* = 357 (195 INT; 162 CONT) from 2 intervention classes and 1 control classes in 1 school in two urban and one rural in Ireland. Sex F/M (n): 173/184 Age: 6-11	Total duration: 26 weeks Frequency: 2 times per week Length: 25 minutes	Project Energize program: a primary school-based PA, health, and nutrition intervention program. Each of the PE lessons incorporated the revision of four FMS activities that accounted for cultural, environmental and curriculum differences.	PE teachers	TGMD-2	Measure: Not reported Delivered as intended: Not reported	The INT group's locomotor standard score, object-control standard score, and GMQ significantly improved from pre- to post-intervention compared to the CONT group, which significantly dis-improved. A group-time interaction effect was found in favor of the INT group for locomotor standard score, object-control standard score, and GMQ.

Burns et al. (2017a) Effect	QE	PE, DS, SI	*N* = 1,460 from 3 schools in USA. Sex F/M (n): 730/730 Age: 8.4 ± 1.8	Total duration:12 weeks Frequency: one, plus recess and classroom PA (once a week required suggested 3 times a week) -Length: 50 minutes for PE, 10 for classroom PA minutes	CSPAP on gross motor skills development: Dynamic Physical Education for Elementary School Children curriculum during PE lessons, PA engagement opportunities throughout the school day during recess and regular classroom time (during which teachers integrated PA into academic lessons and classroom activity breaks via stretching, walking, jumping, or relaxation activities), and SI that provided teacher professional training to increase the quality of PE.	PE teachers and classroom teachers	TGMD-2	Measure: Observations (field notes) and teacher reports - Delivered as intended: High	The students' MC percent scores improved at the post-test compared with the baseline (p<.001).
Burns et al. (2017b) School	QE	PE, PADS, PABAS,SI	*N* = 976 from 5 schools in Salt Lake City district in USA. Sex F/M (n): 413/563 Age: 9.1 ± 1.5	Total duration: 36 weeks Frequency: PE= 1 day, Recess 2 a day (10 a week), MI= at least one per day Length: PE= 50 min, Recess=15 min, MI= 10 min	CSPAP program: Aimed to provide children PA opportunities throughout the school day during specific leisure times (i.e., recess) and integrated PA that focused on motor skill development into academic lessons and classroom activities with continuous teacher training and assistance.	PE teachers, physical activity leaders, and classroom teachers	TGMD-3	Measure: confirmed by interviews and spot checks by research team Delivered as intended: High	There were statistically significant coefficients for time (p<.001) and an age × time interaction (p<.001) on children's MC. Significant improvements were reported for locomotor and object control skills.
Burrows et al. (2014)	QE	PABAS	*N* = 40 (25 games- program group, 15 sports-program group) in rural area in Canada Sex F/M (n): 16/9 (games), 7/8 (sports) Age:	Total duration: 11 weeks Frequency: 5 times per week Length: 60 minutes	Two after school programs: A low-organized games (outside) and sports-based program (indoors) model including swimming activities, floor hockey, and soccer activities.	After school program leaders	TGMD-2	Measure: Not reported Delivered as intended: Not reported	The sports-based program participants showed no improvement in FMS (p=.91), and the games-based program participants significantly improved their proficiency (p<.05).

			7.87 ± 1.07 (games) 8.37 ± 1.27 (sports)						
Chan et al. (2019)	CRCT	PE	*N* = 282 from 5 intervention classes and 5 control classes in 5 schools in urban in China Sex F/M (n): 194/84 Age: 8.4 ± 0.56	Total duration: 13 weeks (a total 550 minutes) Frequency: 1-2 times per week Length: 45-70 minutes	Assessment for learning intervention program: Fun movement activities related to specific motor skills were implemented and aligned with the evaluation criteria of TGMD-3 to facilitate formative assessment.	PE teachers	TGMD-3	Measure: Observations of program delivery using checklist, lesson plans Delivered as intended: High	The INT group's locomotor skills and overall FMS competence from baseline to post-intervention were greater (p<.001) than the CONT group. There was no significant difference between the groups for object control skills (p=.116) .
Chen et al. (2016)	QE	PE	*N* = 4,234 (1377 soccer; 1496 overhand; 1361 forehand) from 9 schools in suburban in USA Sex F/M (n): 2,032/2,202 Age: 6-11	Total duration: 2 years Frequency: 1 time per week Length: 60 minutes	CATCH PE Curriculum: designed as a comprehensive elementary school PE and nutrition program, which includes motor skill themes and physical fitness activities.	PE teachers	PE Metrics	Measure: Video recorded assessed using Quality of Teaching Rubrics, lesson plans Delivered as intended: High	The CATCH PE was conducive to improving fourth- and fifth-grade students' motor skill competency in the three FMS skills.
Cliff et al. (2007)	QE	PABAS,FCE	*N*= 13 in Australia Sex F/M (n): 8/5	Total duration: 10 weeks Frequency:	SHARK program: a community-based program with an additional home-based PA motor development program.	Researcher and parents	KTK	Measure: Not reported	There were significant improvements from pre- to post-test in

			Age: 10.4 ± 1.2	1 time per week Length: 120 minutes	Family participation and parental support along with skill practice outside the program (home tasks/activities). Researched approach centered around improving motor skills and social support in children to increase their PA, self-esteem, and enjoyment.			Delivered as intended: Not reported	the intervention (p<.001).
Cohen et al. (2015)	CRCT	PE, PADS, SI, FCE	*N* = 460 (199 INT; 261 CONT) from 25 classes in 8 schools in Australia Sex F/M (n): 173/184 Age: 8.5 ± 0.7	Total duration: 48 weeks Frequency: 1 time per week Length: 120 minutes	SCORES: A multicomponent PA and FMS intervention program through teacher professional learning and student leadership workshops Used a range of approaches targeting the home environment (newsletters, parent evening, and FMS homework) to engage parents and encourage them to support their children's MC.	Teachers, principals, parents, and community leaders	TGMD-2	Measure: Observation checklist Delivered as intended: High	There was a statistically significant group–time interaction for overall FMS, with children in the intervention group scoring significantly higher (p=.045) than those in the CONT group.
Gu et al. (2017)	QE	PE	*N* = 273 (200 INT; 73 CONT) from 3 schools in USA Sex F/M (n): 137/136 Age: 10.89 ± 0.80	Total duration: 8 weeks Frequency: 3 times per week Length: 45 minutes	Goal setting and Pedometer based intervention: A goal-setting strategy stimulated elementary school children's adaptive motivation during PE lessons by encouraging them to reach their goals through the use of pedometers.	PE teachers	PE Metrics	Measure: Not reported Delivered as intended: Not reported	The INT group had significantly higher MC compared to the CONT group (p<.001).
Johnstone et al. (2017)	QE	PADS	*N* = 336 (291 INT; 45 CONT) from 11 classes in 7 schools in UK	Total duration: 20 weeks Frequency: 2 times per week	Go2Play Active Play - the program combined structured games and free play to increase children's FMS.	Research team	TGMD-2	Measure: Not reported Delivered as	The INT group had significant interaction for GMQ score (p<.001) and percentile (p = 0.04), locomotor skills

			Sex F/M (n): 152/139 (INT), 30/15 (CONT) Age: 7 ± 1.1 (INT), 7.4 ± 0.9 (CONT)	Length: 60 minutes				intended: Not reported	score and percentile (both p=0.02) but no significant interaction for object control skills score (p=0.1) and percentile (p=0.3) at pre- and post-assessment.
Lamml e et al. (2016)	CRCT	PADS	*N* = 1,736 (957 INT; 779 CONT) from 81 intervention classes and 76 control classes in 91 schools in Germany Sex F/M (n): 866/870 Age: 7.1 ± 0.6	Total duration: One school year Frequency: daily from 13 PA units Length: 10-15 minutes	Join the Healthy Boat: A school-based, teacher-centered health promotion program. The INT program includes 13 PA teaching units and short daily exercises in class to increase PA, which work to target an increase in children's motor skills.	PE teachers and trained research Staff	DKT	Measure: Not reported Delivered as intended: Not reported	The INT group showed significant improvement in the motor skills (p<.05) in comparison to the CONT group.
Laurent et al. (2018)	CRCT	PABAS, FCE	*N* = 28 (17 INT; 11 CONT) from 1 school in USA Sex F/M (n): 15/13 Age: 9.3 ± 1.5	-Total duration: 6 weeks -Frequency: 2 times per week -Length: 60 minutes	Resistance-training program, movement programs from university or off-campus training facility.	Coaches and research team	Functional Movement	Measure: Observation checklist Delivered as intended: High	The INT group achieved greater improvements in FMS score (p<.001), relative to CONT group.
Laukka nen et al. (2015)	RCT	FCE	*N* = 89 (44 INT; 45 CONT) from 2 intervention classes and 1 control classes in 1 school in Finland	Total duration: 48 weeks Frequency: N/A Length: N/A	The program took place at a university or other off-school site (community-based). Family involvement and counseling parents lead to an increase in children's MC.	Parents and research team	KTK and TCB	Measure: Observation checklist Delivered as	The INT group's mean score of KTK (p<.001) increased significantly with time, not TCB.

			Sex F/M (n): 173/184 Age: 6.16 ± 1.13					intended: High	
Lee at al. (2020a)	QE	PABAS	*N* = 31 (20 INT; 11 CONT) from 1 school in USA Sex F/M (n): 18/13 Age: 6.65 ± 0.91	Total duration: 8 weeks Frequency: 3 times per week Length: 60 minutes	The FMS INT program aimed to promote FMS competence by focusing on the mastery of 12 basic motor skills.	Research team	TGMD-2	Measure: Field observations Delivered as intended: High	The INT group who participated in the FMS-based afterschool program showed significant improvements in FMS competence, compared to the CONT group.
Lee et al. (2020b)	QE	PABAS	*N* = 35 (24 INT; 11 CONT) from 1 intervention classes and 2 control classes in 3 schools in USA Sex F/M (n): 22/13 Age: 6.52 ± 0.97	Total duration: 8 weeks Frequency: 3 times per week Length: 60 minutes	INT program aimed at teaching the 12 basic motor kills from the TGMD-2 criteria. CONT group program included unsupervised free play and academic tutoring with no motor skill related instructions.	Research team	TGMD-2	Measure: Not reported Delivered as intended: Not reported	There were significant group differences between the INT and CONT group in FMS competence ($p<.001$), non-significant gender differences between boys and girls in FMS competence ($p=.85$), and non-significant interaction effects over time ($p=.52$). The effect size reported significant improvements on the INT group's FMS competence, with a medium to large effect size (range gs 0.49–1.92), in comparison to the CONT group.

Maskell et al. (2004)	CRCT	PE	*N* = 42 (20 INT; 22 CONT) from 1 intervention class and 1 control class in 1 school in urban in USA Sex F/M (n): 23/19 Age: 6.98 ± 0.42	Total duration: 5 weeks Frequency: N/A (total 16 lessons) Length: 30 minutes	Brain Gym movement program: Designed to engage students in moving different body parts across the body midline. The program is a series of simple-challenging fundamental movement skills intended to enhance cognitive processing, psychomotor and whole-brain learning.	Research team	TGMD-2	Measure: Not reported Delivered as intended: Not reported	There were no significant pre- or post- test group differences in TGMD-2 scores (p =.26).
McKenzie et al. (1998)	CRCT	PE, SI	*N* = 467 (201 INT; 266 CONT) from 56 classes in 7 schools in suburban area in USA Sex F/M (n): 109/358 Age: N/A	Total duration: 2 years Frequency: 3 times per week Length: 30 minutes	SPARK PE curriculum: A comprehensive program designed to enhance both children's health related physical fitness and sports/motor skills as they participated in high levels of enjoyable PA through a classroom teacher professional development program.	Trained PE teachers and classroom teachers	N/A	Measure: Not reported Delivered as intended: Not reported	The INT group reported the most improvement in total motor skills—gain scores were significant for catching (p=.005) and throwing (p =.008). INT effects did not differ by gender or grade.
McWhannell et al. (2018)	CRCT	PABAS	*N* = 146 from 16 schools in Northern Ireland Sex F/M (n): 89/57 Age: 9.6 ± 0.3	Total duration: 26 weeks Frequency: 2 times per week Length: 60 minutes	A-CLASS Project: Aimed to quantify the effectiveness of structured and unstructured PA programs on children's PA, FMS, physical self-perception, and self-esteem.	Trained coaches	N/A	Measure: Not reported Delivered as intended: Not reported	Boys in the INT group were significantly more proficient than girls, being more improved in seven out of the eight motor skills; however these differences were only significant (p<.01) in the four object control skills (catch, overarm throw, strike, and kick).

Miller et al. (2016)	RCT	PE, SI	*N* = 107 (55 INT; 52 CONT) from 4 classes in 1 school in Australia Sex F/M (n): 48/59 Age: 10.87 ± 0.87	Total duration: 6 weeks Frequency: 1 time per week Length: 60 minutes	PLUNGE: PE lessons based on game-centered curriculum. Aimed to improve the children's FMS, game skills, in-class MVPA and enjoyment of PA through a professional learning process involving classroom teacher education and mentoring.	Classroom teachers	TGMD-2	Measure: lesson observations Delivered as intended: High	The INT group's treatment effect significantly increased in FMS (throw and catch; p<.001). The treatment effect significantly increased the INT group's FMS (p<.001).
Miller et al. (2015)	CRCT	PE	*N* = 168 (97 INT; 71 CONT) from 6 classes in 7 schools in Australia Sex F/M (n): 110/58 Age: 11.2 ± 1.00	Total duration: 7 weeks Frequency: N/A Length: N/A	PLUNGE program: Aimed to increase the complexity of challenges experienced through gameplay-situated learning for the improvement of FMS.	Researchers and teachers	TGMD-2	Measure: observations Delivered as intended: High	The treatment effect significantly increased the INT group's FMS (p<.001).
Nathan et al. (2017)	QE	PE, PADS, SI	*N* = 174 (83 INT; 91 CONT) from 2 schools in Australia Sex F/M (n): 89/85 Age: 6.1 ± 0.9	Total duration: 10 weeks Frequency: 2 times per week Length: 30 minutes	The GLASS program: included trained students who instructed their peers to improve FMS during PE lessons and classroom settings, and trained teachers supporting their peers' instruction, which contributed an SI component to the program.	Classroom teachers and peer leaders (students)	TGMD-3	Measure: checklist, observation Delivered as intended: Not reported	The treatment effect on the INT group's overall object control skills was statistically significant (p<.001).
Okely et al. (2017)	QE	PADS	*N* = 407 (223 INT; 184 CONT) from 8 schools in Australia	Total duration: 44 months Frequency:	PALDC program: A whole-of-school health promotion approach.	Classroom teachers	N/A	Measure: Not reported	There was a significantly greater increase in the INT group's total FMS (p<.01),

			Sex F/M (n): 307/100 Age: 8.69 ± 1.7	2 times per week Length: 25 minutes	Aimed to develop children's FMS with a focus on initiating sustainable changes in the delivery of FMS in a school context by modifying the physical and social environment and developing links with the home and local community.			Delivered as intended: Not reported	compared with the CONT group.
Platvoet et al. (2016)	QE	PE	*N* = 425 (294 INT; 131 CONT) from 6 schools in Netherlands Sex F/M (n): 233/192 Age: 6.4 ± 0.52	Total duration: 4 weeks Frequency: 2 times per week Length: 45 minutes	A goal-directed learning program: Aimed to improve FMS performance based on information in the Dutch national handbook for PE in primary education. PE teachers stimulated goal-directed learning, skill-specific exercises, and individual practice.	PE teachers	KTK	Measure: Not reported Delivered as intended: Not reported	The INT group improved their FMS performance significantly more than the CONT group.
Rudd et al. (2016)	QE	PE	*N* = 333 (135 INT; 198 CONT) from 14 intervention classes and 6 control classes in 3 schools in Australia Sex F/M (n): 169/164 Age: 8.1±1.1	Total duration: 16 weeks Frequency: 1 time per week Length: 120 minutes	Gymnastics curriculum developed by Gymnastics Australia: Aimed to develop stability, locomotive and object control skills, and general body coordination.	Gymnastics coaches, classroom teachers, and PE teachers	TGMD-2 and KTK	Measure: Lesson observations Delivered as intended: High	The INT group showed a significant improvement compared to the CONT group in stability and object control ($p<.05$) but not in locomotor skills ($p>.05$).
Rudd et al. (2017)	QE	PE	*N* = 113 (56 INT; 57 CONT) from 2 classes in 1 school in urban area in Australia	Total duration: 8 weeks Frequency: 2 times per week	LaunchPad program: gymnastics curriculum aimed to develop children's MC.	PE teachers	KTK and TGMD-2	Measure: Not reported Delivered as	The INT group showed a significant improvement compared to the CONT group in FMS.

			Sex F/M (n): 52/61 Age: 9.4±1.8	Length: 60 minutes				intended: Not reported	
Silveira et al. (2018)	QE	PE, FCE, SI	*N* = 357 (17 INT; 18 CONT) from 2 intervention classes and 1 control classes in 1 school in Brazil Sex F/M (n): 173/184 Age: 7-12	Total duration: 12 weeks Frequency: 2 times per week Length: 60 minutes	ACTION FOR HEALTH program: a multicomponent intervention program included physical exercise sessions during PE lessons (e.g., circuit training, aerobic/sports activities, and recreational games, etc.), parent support to promote PA during after school classes, and nutritional education sessions (e.g., goal setting and dietary counselling with parents).	PE teachers, medical or healthcare staff, parents, and research team	KTK	Measure: Not reported Delivered as intended: Not reported	The INT group reported significant interaction (group*time) improvements (p<.001).
Skowroński et al. (2019)	QE	PE, PABAS	*N* = 31 (20 INT; 11 CONT) from 2 classes in 2 schools in urban area in Poland Sex F/M (n): 16/15 Age: 7.19 ± 0.28	Total duration: 1 school year (1440 min total) Frequency: 4 times per week Length: 45 minutes	From Fun to Sport program: PE lessons combined with an extracurricular after school program with an emphasis on the development of children's FMS	PE teachers	TGMD-2	Measure: Not reported Delivered as intended: Not reported	There were statistically significant differences in the level of FMS between the INT group and CONT group at post-test (p<.05).

Notes: A-CLASS=Active City of Liverpool, Active Schools and SportsLinx; CATCH=Child and Adolescent Trial for Cardiovascular Health; CONT=Control; CRCT=Cluster Randomized Control Trial; DTK=Dordel Kock Test; FCE=Family and Community Engagement; FMS=Fundamental Motor Skill; GLASS=Great Leaders Active StudentS; GMQ=Gross Motor Quotient; INT=Intervention; KTK=Körperkoordinationstest Für Kinder; MC=Motor Competence; MVPA=Moderate to Vigorous Physical Activity; PA=Physical Activity; PABAS=Physical Activity Before and After School; PADS=Physical Activity During School; PALDC=Physical Activity in Linguistically Diverse Communities; PE=Physical Education; PLUNGE=Professional Learning for Understanding Games Education; QE=Quasi Experimental; RCT= Randomized Control Trial; SCORES=Supportive Children's Outcome using Rewards, Exercise, and Skills; SI=Staff Involvement; SPARK=Sports, Play and Active Recreation for Kids; TCB=Throwing and Catching a Ball; MI=Movement integration; TGMD=Test of Gross Motor Development.

APPENDIX D

THE PRISMA 2020 MAIN CHECKLIST FOR STUDY 2

Table D.1 The PRISMA 2020 Checklist for Study 2

Section and Topic	Item #	Checklist item	Location where item is reported
TITLE			
Title	1	Identify the report as a systematic review.	53
ABSTRACT			
Abstract	2	See the PRISMA 2020 for Abstracts checklist.	N/A
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of existing knowledge.	53-56
Objectives	4	Provide an explicit statement of the objective(s) or question(s) the review addresses.	55-56
METHODS			
Eligibility criteria	5	Specify the inclusion and exclusion criteria for the review and how studies were grouped for the syntheses.	56-57
Information sources	6	Specify all databases, registers, websites, organisations, reference lists and other sources searched or consulted to identify studies. Specify the date when each source was last searched or consulted.	57-58
Search strategy	7	Present the full search strategies for all databases, registers and websites, including any filters and limits used.	58-59
Selection process	8	Specify the methods used to decide whether a study met the inclusion criteria of the review, including how many reviewers screened each record and each report retrieved, whether they worked independently, and if applicable, details of automation tools used in the process.	57-59
Data collection process	9	Specify the methods used to collect data from reports, including how many reviewers collected data from each report, whether they worked independently, any processes for obtaining or confirming data from study	57-59

Section and Topic	Item #	Checklist item	Location where item is reported
		investigators, and if applicable, details of automation tools used in the process.	
Data items	10a	List and define all outcomes for which data were sought. Specify whether all results that were compatible with each outcome domain in each study were sought (e.g. for all measures, time points, analyses), and if not, the methods used to decide which results to collect.	N/A
	10b	List and define all other variables for which data were sought (e.g. participant and intervention characteristics, funding sources). Describe any assumptions made about any missing or unclear information.	58
Study risk of bias assessment	11	Specify the methods used to assess risk of bias in the included studies, including details of the tool(s) used, how many reviewers assessed each study and whether they worked independently, and if applicable, details of automation tools used in the process.	61-62
Effect measures	12	Specify for each outcome the effect measure(s) (e.g. risk ratio, mean difference) used in the synthesis or presentation of results.	59-60
Synthesis methods	13a	Describe the processes used to decide which studies were eligible for each synthesis (e.g. tabulating the study intervention characteristics and comparing against the planned groups for each synthesis (item #5)).	58-61
	13b	Describe any methods required to prepare the data for presentation or synthesis, such as handling of missing summary statistics, or data conversions.	58-61
	13c	Describe any methods used to tabulate or visually display results of individual studies and syntheses.	91
	13d	Describe any methods used to synthesize results and provide a rationale for the choice(s). If meta-analysis was performed, describe the model(s), method(s) to identify the presence and extent of statistical heterogeneity, and software package(s) used.	60-61
	13e	Describe any methods used to explore possible causes of heterogeneity among study results (e.g. subgroup analysis, meta-regression).	60-61
	13f	Describe any sensitivity analyses conducted to assess robustness of the synthesized results.	N/A

Section and Topic	Item #	Checklist item	Location where item is reported
Reporting bias assessment	14	Describe any methods used to assess risk of bias due to missing results in a synthesis (arising from reporting biases).	61
Certainty assessment	15	Describe any methods used to assess certainty (or confidence) in the body of evidence for an outcome.	N/A
RESULTS			
Study selection	16a	Describe the results of the search and selection process, from the number of records identified in the search to the number of studies included in the review, ideally using a flow diagram.	62, 91
	16b	Cite studies that might appear to meet the inclusion criteria, but which were excluded, and explain why they were excluded.	91
Study characteristics	17	Cite each included study and present its characteristics.	62-71, 191-207
Risk of bias in studies	18	Present assessments of risk of bias for each included study.	92-93
Results of individual studies	19	For all outcomes, present, for each study: (a) summary statistics for each group (where appropriate) and (b) an effect estimate and its precision (e.g. confidence/credible interval), ideally using structured tables or plots.	62-71
Results of syntheses	20a	For each synthesis, briefly summarise the characteristics and risk of bias among contributing studies.	62
	20b	Present results of all statistical syntheses conducted. If meta-analysis was done, present for each the summary estimate and its precision (e.g. confidence/credible interval) and measures of statistical heterogeneity. If comparing groups, describe the direction of the effect.	71-76
	20c	Present results of all investigations of possible causes of heterogeneity among study results.	82-83
	20d	Present results of all sensitivity analyses conducted to assess the robustness of the synthesized results.	N/A
Reporting biases	21	Present assessments of risk of bias due to missing results (arising from reporting biases) for each synthesis assessed.	63-71
Certainty of evidence	22	Present assessments of certainty (or confidence) in the body of evidence for each outcome assessed.	N/A
DISCUSSION			

Section and Topic	Item #	Checklist item	Location where item is reported
Discussion	23a	Provide a general interpretation of the results in the context of other evidence.	76-88
	23b	Discuss any limitations of the evidence included in the review.	81-82
	23c	Discuss any limitations of the review processes used.	88-89
	23d	Discuss implications of the results for practice, policy, and future research.	89-90
OTHER INFORMATION			
Registration and protocol	24a	Provide registration information for the review, including register name and registration number, or state that the review was not registered.	56
	24b	Indicate where the review protocol can be accessed, or state that a protocol was not prepared.	56
	24c	Describe and explain any amendments to information provided at registration or in the protocol.	56
Support	25	Describe sources of financial or non-financial support for the review, and the role of the funders or sponsors in the review.	N/A
Competing interests	26	Declare any competing interests of review authors.	N/A
Availability of data, code and other materials	27	Report which of the following are publicly available and where they can be found: template data collection forms; data extracted from included studies; data used for all analyses; analytic code; any other materials used in the review.	189-207

APPENDIX E

ARTICLE SEARCH STRATEGY AND KEY TERMS FOR STUDY 2

SOCIAL AND EMOTIONAL LEARNING OUTCOMES

Table E.1 Search Strategy and Key Terms for Study 2

Population	Study Design	Intervention Focus	Social and emotional learning related outcomes	Database
Child*[tiab] OR Child*[mh] OR OR Boys[tiab] OR Girls[tiab] OR Student[tiab] OR Student*[tiab]	Intervention[tiab] OR Experiment[tiab] OR Program*[tiab] OR Evaluation[tiab] OR Trial[tiab] OR Random*[tiab] OR Clinic*[tiab] OR “Controlled trial”[tiab]	“Physical activity” [tiab] OR Exercise[tiab]	(Social[TIAB] OR Emotional[TIAB] OR Socioemotional[TIAB] OR “Social and emotional”[TIAB] OR Affect*[TIAB] OR Interpersonal[TIAB] OR “Positive youth development”[TIAB] OR Coping[TIAB] OR “Self-esteem”[TIAB] OR Prosocial[TIAB] OR “Self-awareness”[TIAB] OR “Self-management”[TIAB] OR “Self-control”[TIAB] OR Responsib*[TIAB] OR Relationship[TIAB] OR Mindfulness[TIAB] OR Assertiveness[TIAB] OR “Valuing diversity”[TIAB] OR “Perspective taking”[TIAB] OR “Self-talk”[TIAB] OR “Identifying emotions”[TIAB] OR “Problem solving”[TIAB] OR “Self-regulation”[TIAB] OR Resilience[TIAB] OR “Emotion expression”[TIAB] OR “Challenging behavior”[TIAB] OR "Problem behavior”[TIAB] OR “Conflict resolution”[TIAB] OR “Decision-making”[TIAB] OR “Inhibitory control”[TIAB] OR Cooperation[TIAB] OR Prosocial[TIAB] OR Internalizing[TIAB] OR Externalizing[TIAB] OR On-task[TIAB])	PubMed/ Medline

Youth* or Child* or Child* or Student* or Boys or Girls or teenager* or teen*	Intervention or Experiment or Program or Evaluation or Trial or Random* or Clinic* or Controlled trial	Physical activity or Exercise	Social or Emotional or Socioemotional or Social and emotional or Affect* or Interpersonal or Positive youth development or Coping or Self-esteem or Prosocial or Self-awareness or Self-management or Self-control or Responsib* or Relationship or Mindfulness or Assertiveness or Valuing diversity or Perspective taking or Self-talk or Identifying emotions or Problem solving or Self-regulation or Resilience or Emotion expression or Challenging behavior or Problem behavior or Conflict resolution or Decision-making or Inhibitory control or Cooperation or Prosocial or Internalizing or Externalizing or On-task	PsycINFO SPORTDiscus ERIC CINAHL
(youth*:ab,ti OR child*:ab,ti OR student*:ab,ti OR boys:ab,ti OR girls:ab,ti) AND (intervention:ab,ti OR experiment:ab,ti OR program*:ab,ti OR evaluation:ab,ti OR trial:ab,ti OR random*:ab,ti OR clinic*:ab,ti OR 'controlled trial':ab,ti) AND ('physical activity':ab,ti OR exercise:ab,ti) AND (social:ab,ti OR emotional:ab,ti OR socioemotional:ab,ti OR 'social and emotional':ab,ti OR affect*:ab,ti OR interpersonal:ab,ti OR 'positive youth development':ab,ti OR coping:ab,ti OR 'self-esteem':ab,ti OR 'self-awareness':ab,ti OR 'self-management':ab,ti OR 'self-control':ab,ti OR responsib*:ab,ti OR relationship:ab,ti OR mindfulness:ab,ti OR assertiveness:ab,ti OR 'valuing diversity':ab,ti OR 'perspective taking':ab,ti OR 'self-talk':ab,ti OR 'identifying emotions':ab,ti OR 'problem solving':ab,ti OR 'self-regulation':ab,ti OR resilience:ab,ti OR 'emotion expression':ab,ti OR 'challenging behavior':ab,ti OR 'problem behavior':ab,ti OR 'conflict resolution':ab,ti OR 'decision-making':ab,ti OR 'inhibitory control':ab,ti OR cooperation:ab,ti OR prosocial:ab,ti OR internalizing:ab,ti OR externalizing:ab,ti OR 'on task':ab,ti)				Embase

APPENDIX F

OVERVIEW OF CHARACTERISTICS OF INCLUDED STUDIES EXAMINING SOCIAL AND EMOTIONAL LEARNING

Table F.1 Characteristics of Included Studies for Study 2

Study	Study Design	CSPAP Component	Population, Sample size, Sex, and Age	Description of Intervention	Control Group	Relevant SEL Outcome(s)	Main Findings in SEL Outcomes as Reported	Fidelity Measure
Andrade et al. (2020)	CRCT	PE	*N* = 140 (68 INT; 72 CONT) from 5 intervention classes and 5 control classes in 1 school in a local urban area in Brazil Sex F/M (n): 81/59 Age: 7-11	Name: ExerGame Total duration: 2 weeks Frequency: 3 times per week Length: 40 minutes Design: Teacher-driven Training Support: by research team	PE as usual	Tension Depression Anger Vigor Fatigue Mental confusion Self-esteem	The main results of the INT group demonstrated reduced tension in girls (p<.05; ES: 0.62). Anger was lower for girls (p<.05; ES: 0.61) in the INT group. Vigor was higher for girls in the INT group than for in those in the PE group (p<.05; ES: 0.56). The main results of the INT group indicated increased self-esteem in boys (p<.05; ES: 0.58) and reduction of girls' mental confusion (p<.05; ES 0.58).	Not reported
Annesi et al. (2017)	CRCT	PABAS	*N* = 141 (86 INT; 55 CONT) in the southeastern USA Sex F/M (n): 63/78 M_{age}: 10	Name: Youth Fit 4 Life program Total duration: 12 weeks Frequency:	Typical in the allotted physical activity time as usual during	Exercise self-efficacy Negative mood Self-regulation for physical activity	The INT program was associated with significantly greater improvements in self□regulation, mood, self□efficacy over 3 months. Changes in self□regulation, mood and self□efficacy significantly mediated the	Survey and structured fidelity checks

				4 times per week Length: 45 minutes Design: Researchers (after-school care counselors)-driven Training Support: 6-hour training by researchers	after□ school care		treatment type–BMI relationship over 3 months (R^2 =.12, p=.002), with change in self□regulation being a significant independent mediator. Changes in BMI and self□regulation reciprocally reinforced one another. Gender was not a significant moderator of those relationships.	
Annesi et al. (2016 a)	RCT	PABAS	*N* = 114 (72 INT; 42 CONT) in the southeastern USA Sex F/M (n): 63/53 Age: 5-8	Name: Youth Fit 4 Life program Total duration: 24 weeks Frequency: 4 times per week Length: 45 minutes Design: Researchers-driven Training Support: 6-hour training by researchers	Typical in the allotted physical activity time as usual during after□ school care	Exercise barriers Self-efficacy Physical self-concept	Improvements in physiological measures and exercise self-efficacy were significantly greater in the INT group. The INT group demonstrated significantly greater improvements in the physiological measures, with its treatment's theoretical basis and application within after-school care both supported.	Survey and structured fidelity checks
Annesi et al. (2016 b)	RCT	PABAS	*N* = 145 (88 INT; 57 CONT) in the USA Sex F/M (n): 65/80 Age: 9-12	Name: Youth Fit 4 Life program Total duration: 12 weeks Frequency: 4 times per week Length: 45 minutes Design: Researchers (after-school care counselors)-driven	Typical in the allotted physical activity time as usual during after□ school care	Self-regulation Negative mood Exercise self-efficacy	Increases over 12 weeks in out-of-school physical activity and improvements in self-regulation for physical activity, exercise self-efficacy and mood were significantly greater in the INT group when contrasted with a CONT group. Changes in the psychosocial variables significantly mediated the group–physical activity change relationship (R^2 = .31, p<.001). Change in self-regulation was a significant independent mediator and had a reciprocal relationship with change in out-of-school physical activity.	Survey and structured fidelity checks

				Training Support: 6-hour training by researchers				
Brown & Fry (2014)	QE	PABAS	*N* = 48 from 1 school in the USA Sex F/M (n): 48/0 Age: 8-10	Name: Strong Girls Total duration: 15 weeks Frequency: 1 time per week Length: 120 minutes Design: Researchers-driven Training Support: 2-hour training by researchers	No control group	Task-involving Ego-involving Caring Social Behavior Global Task orientation Ego orientation	Perceptions of the caring climate were positively and significantly correlated with global self-worth, $r(24) = .53$, $p<.01$. Perceptions of an ego-involving climate were negatively and significantly correlated with academic competence, $r(21) = -.44$, $p<.05$. The only significant difference in the scores was in task-orientation pre- ($M_{pre\text{-}task}$ = 4.45, SD = .56) and post-Strong Girls ($M_{post\text{-}task}$ = 4.61, *SD* = .73); $t(26) = -2.16$, $p=.04$.	Not reported
Bundy et al. (2017)	CRCT	PADS, PABAS	*N* = 221 (113 INT; 108 CONT) from 6 INT classes and 6 CONT classes in 12 schools in an urban area in Australia Sex F/M (n): 102/119 Age: 5-7	Name: The Sydney Playground INT program Total duration: 13 weeks Frequency: Not reported Length: Not reported Design: Researchers-driven Training Support: 2-hour training for teachers and parents by researchers	Recess as usual	Psychosocial skills and behaviors: communication, cooperation, assertion, responsibility, empathy, engagement, and self□control Problem behaviors: externalizing, bullying, hyperactivity/inattention, internalizing, and autism spectrum Self-cognitive Self-peer Self-physical Self-maternal	Play and Social Interaction: Increases in observed time spent in play in the INT group did not reach statistical significance (p=.08). Measurable differences in observations of the number of playmates were not altered by the INT (p=.31). There were no changes from the INT (p= .11□.6); similarly, most baseline scores on social skills fell within the expected range. No effect of the INT was found on SSIS□RS scores (p= .1□.4).	Process evaluation (observations and interviews, videos and photographs, record of how children were playing, and how equipment stored and use)
Buišić & Đorđić (2018)	QE	PE	*N* = 100 (45 INT; 55 CONT) from 4 classes in 2 schools in Serbia Sex F/M (n): 46/54 Grade: 3rd	Name: Not reported Total duration: 8 weeks (24 lessons) Frequency: Not reported	PE as usual	Personal responsibility Social responsibility Amotivation External regulation Introjected regulation Identified regulation Intrinsic motivation	After completing the INT program, statistically significant differences were detected in all the dependent variables, consistently in favor of the INT group TPSR model produced the largest effects in the domain of responsibility and self-determined motivation among the	Not reported

				Length: 45 minutes Design: Researchers-driven Training Support: By researchers		Social acceptance Behavioral conduct Classmate support	students, which points to the validity of its implementation in regular PE teaching.	
Chan et al. (2019)	CRCT	PE	*N* = 276 (147 INT; 129 CONT) from 5 INT classes and 5 CONT classes in 3 schools in Hongkong Sex F/M (n): Not reported M_{age}: 8.4	Name: A + FMS intervention Total duration: 13 weeks Frequency: 1-2 time(s) per week Length: 45-70 minutes Design: Teachers-driven Training Support: 6-hour training by researchers	PE as usual	Perceived teacher support Enjoyment of PE	The school-based A+ FMS INT program was effective in increasing children's FMS proficiency and perceptions of teacher support, but not their perceived competence or enjoyment of PE.	Process evolution, observation checklist, survey, lesson plan, assessment record sheets and a mid-program meeting
Cocca et al. (2020)	QE	PE	*N* = 229 (102 INT; 127 CONT) from 4 schools in an urban area in Mexico Sex F/M (n): 104/125 Age: 10-12	Name: Not reported Total duration: 1 academic semester Frequency: 2 weeks Length: 45 minutes Design: Researchers-driven Training Support: By researchers	PE as usual	Self-esteem Psychological wellbeing Anxiety Stress	Both INT and CONT group reported positive psychological conditions at pre-test, which did not change after the INT. Participants' psychological health was already excellent at pre-test, as they reported very high psychological well-being and self-esteem at the same time, they scored very low in anxiety and stress.	Not reported
Coolkens et al. (2018)	CRCT	PE, PADS	*N* = 281 (136 supervised, 136 organized, 9 excluded) from 7 supervised recess groups in 7 schools	Name: Parkour Unit Total duration: 20 weeks Frequency:	No control group; two different conditions (supervise	Social behavior	During supervised recess, girls spent more time (4% vs 2%, p= .02) in prosocial behavior than boys.	Observation

			and 7 organized recess groups in 7 schools in Belgium Sex F/M (n): 144/137 Age: 8-10	1-2 time(s) per week Length: 50 minutes PE and 20 minutes recess option Design: Researchers-driven Training Support: 3-hour training by researchers	d recess groups and organized recess groups)			
Duck et al. (2020)	QE	PADS, FCE	*N* = 35 (17 INT; 18 CONT) from 1 INT class and 1 CONT class in 1 school in a rural area in the USA Sex F/M (n): 23/12 Age: 9.20±0.41	Name: UNICEF Power Up Curriculum Program Total duration: 10 weeks Frequency: Not reported Length: Not reported Design: Teachers-driven Training Support: By research team	Waitlisted	Prosocial behavior (altruism)	The baseline prosocial behavior score did not differ by group at baseline (p= .13) or post-intervention (p= .73). The prosocial scale non-significantly decreased slightly from baseline to post-intervention in both the INT and CONT groups. The INT group decreased from a mean of 8.1 (SD=2.1) to 7.8 (SD=1.5) post-intervention (p=.62), while the CONT group decreased from 8.9 (SD=1.3) at baseline to a mean score of 8.0 (SD=1.9) post-intervention (p= .09).	Not reported
Duncan et al. (2018)	CRCT	PE	*N* = 94 (53 INT; 41 CONT) from 4 classes in 2 schools in central England Sex F/M (n): 45/49 Age: 6-7	Name: Integrated Neuromuscular Training Program Total duration: 10 weeks Frequency: 1 time per week Length: 30-40 minutes Design: Researchers-driven Training Support: Not reported	2 stationary PE classes per week	Physical self-efficacy	A significant group (INT vs. CON) × time (pre vs. post) × gender interaction for physical self-efficacy revealed increased physical self-efficacy pre- to post-INT, compared with CON but only for boys (p=.001). For girls, physical self-efficacy was not significantly different before vs. after the 10-week period for INT and CON groups.	Documentation

Duncan et al. (2009)	RCT	PABAS	*N* = 68 (33 INT; 35 CONT) from 1 school in an urban area in the UK Sex F/M (n): 34/34 Age: 10-11	Name: Not reported Total duration: 6 weeks Frequency: 2 times per week Length: 40 minutes Design: Researchers-driven Training Support: Not reported	Waitlisted	Body image (body esteem)	The result reported that, when compared to the CONT group, participation in 6-week circuit training significantly improved body esteem scores post-INT. The improvement in body esteem scores from pre- to post-INT was greater for girls as compared to boys in the INT group.	Not reported
Elbe et al. (2017)	CRCT	PE, PABAS	*N* = 300 (111 INT [team sports]; 104 INT [individual sports]; 85 CONT.) from 4 schools in a rural area and 4 schools in an urban area in Denmark Sex F/M (n): 158/142 Age: 9.30±0.35	Name: Not reported Total duration: 10 months Frequency: 1 time per week Length: 90 minutes Design: Researchers-driven Training Support: 30-hour training by researchers	PE as usual	Social (team cohesion) Enjoyment of physical activity	Results reported that enjoyment and cohesion (social) measured at the beginning of the INT significantly predicted fitness improvements achieved after 10 months. No differing developmental effects over time could be found in the INT groups with regard to cohesion and enjoyment when compared to the CONT group. Enjoyment and cohesion (social) significantly decreased in the groups that performed individual sports.	Not reported
Escartí et al. (2010)	QE	PE	*N* = 42 (21 INT; 21 CONT) from 1 class in Spain Sex F/M (n): 20/22 Age: 11-12	Name: TPSR PE program Total duration: 1 academic year Frequency: 2 times per week Length: 60 minutes Design: Researchers-driven	PE as usual	Self-regulatory efficacy Social self-efficacy Assertive self-efficacy	The results reported that the TPSR model was an effective teaching instrument that helped teachers to structure lessons and promote the learning of responsibility. A significant increase was observed in the self-regulatory efficacy of INT group participants (ES: large).	Not reported

				Training Support: 30-hour training by research team				
Fu & Burns (2018)	QE	PADS	*N* = 66 (33 INT; 32 CONT) from 1 INT class and 1 CONT class in 1 school in the western USA Sex F/M (n): 36/30 M_{age}: 11.6	Name: Active Video Gaming Total duration: 18 weeks Frequency: 3 times per week Length: 30 minutes Design: Researchers-driven Training Support: By research team	Three 30-minute time windows of unstructured active free play per week, supervised by their classroom teacher and a research staff	Motivation (perceived competence) Enjoyment Self-efficacy Social support Task value Situational motivation Outcome expectancy	There were no group × time interactions found on any of the motivational variables. There were no statistically significant changes in the INT group from pre-test to post-test for perceived competence (p=.68), enjoyment (p=.17), self-efficacy (p=.07), social support (p=.43), task value (p=.37), situational motivation (p=.62), or outcome expectancy (p=.90).	Not reported
Goudas & Magotsiou (2009)	QE	PE	*N* = 114 (57 INT; 57 CONT) from 2 intervention classes and 2 control classes in 4 schools in Greece Sex F/M (n): 55/59 Grade: 6th	Name: Cooperative Program Total Duration: 4 weeks Frequency: 3 times per week Length: 45 minutes Design: Teachers-driven Training Support: By research team	Same subject matter as the INT group, but with a command teaching style	Preference for group learning Preference for individual learning Discomfort in group learning Cooperating skills-self-rating Empathy-self-rating Quick-temperedness-self-rating Disruptiveness self-rating Cooperating skills-peers rating Empathy-peers rating Quick-temperedness-peers rating Disruptiveness-peers rating	The INT group, developed on the basis of specific social skills as learning objectives, reported enhanced social skills and attitudes toward group work shortly after the INT program. The INT group increased their cooperative skills and empathy and decreased their quick-temperedness and their tendency to disrupt compared to those of a CONT group. These findings shown both for self-reports and for peer assessments. The INT group increased their preferences for working in groups and decreased their discomfort with group work.	Not reported
Griffiths & Griffiths (2019)	QE	PE	*N* = 596 (468 INT; 128 CONT) from 4 schools in a local urban area in UK Sex F/M (n): Not reported	Name: Not reported Total duration: Phase 1: 10 days; Phase 2: 5 months	Not reported	Catharsis (to reduce stress or to get away from problems) Social growth	Quantitative analysis reported that the INT program had no impact on attitudes toward social aspects.	Not reported

			Age: 7-12	Frequency: 1 time per week Length: Not reported Design: Researchers-driven Training Support: 1-hour training by research team		(a chance to meet new people) Social continuation (a chance to be with friends) Vertigo (risk with speed, change of position, and location) Aesthetic (involvement in beautiful and graceful movements) Ascetic (sacrificing spare time in order to improve by means of hard and long practices)		
Grillich et al. (2016)	CRCT	PADS	*N* = 925 (432 INT; 493 CONT) from 26 intervention classes and 27 control classes in 45 schools in urban and rural areas in Austria Sex F/M (n): Not reported Age: 8-9	Name: The intervention Classes in Motion Total duration: 1.5 academic years Frequency: Not reported Length: Not reported Design: Teachers-driven Training Support: 20-hour training and 8-hour workshop by research team	PE as usual	Emotional and social experience at school Well-being Attention performance	There were no statistically significant differences between groups for the outcome categories: Emotional and social experience in school, well-being, and attention performance (p = 0.22 to 0.78).	Questionnaire
Hassandra & Goudas (2010)	QE	PE	*N* = 41 (21 INT; 20 CONT) in Greece Sex F/M (n): 19/22 Grade:	Name: Not reported Total duration: 16 weeks Frequency:	Floor hockey unit	Personal responsibility Social responsibility Related knowledge about responsibility	Analysis of quantitative measures did not show significant differences between groups in self-ratings and knowledge of responsibility.	Not reported

			5th	1 time per week Length: 60 minutes Design: Researchers-driven Training Support: Not reported			The qualitative results presented that students developed more articulated perceptions of responsibility.	
Ho et al. (2017)	RCT	PABAS	*N* = 664 (333 INT; 331 CONT) from 12 schools in Hongkong Sex F/M (n): 386/278 M_{age}: 12.3	Name: Positive Youth Development-Based Sports Mentorship Program Total duration: 18 weeks Frequency: 1 time per week Length: 90 minutes Design: Researchers-driven Training Support: 1-day workshop by research team	Exclusive access to a Web-based health education game with 400 questions on healthy lifestyle during the same period	Mental well-being Self-efficacy Resilience Family connectedness School connectedness	The INT improved students' mental well-being (Cohen's d=0.25; p= .001), self-efficacy (Cohen's d=0.22; p=.01), and resilience (Cohen's d=0.19; p=.02).	Observation (random visiting)
Huang et al. (2018)	RCT	PABAS	*N* = 54 (24 INT; 30 CONT) from 2 INT schools and 2 CONT schools in an urban area in the USA Sex F/M (n): 19/35 Age: 9-12	Name: Bicycle Train program Total duration: 4-6 weeks Frequency: 5 times per week Length: Not reported Design: Researchers-driven Training Support: 2–3-hour professional riding safety training by research team	Not reported	Child self-efficacy	The INT group improved in child self-efficacy of 0.84 standard deviation (95% CI [0.37, 1.31]), parent self-efficacy of 0.46 standard deviation (95% CI [0.05, 0.86]), and parent outcome expectations of 0.47 standard deviation (95% CI [0.17, 0.76]) compared to CONT group from Times 1 to 2 (all p<.05).	Not reported

Lwin & Malik (2014)	RCT	PE	*N* = 398 from 12 classes in 3 schools in an urban area in Singapore Sex F/M (n): 165/233 M_{age}: 10.2	Name: Wii exergame Total duration: 6 weeks Frequency: 1 time per week Length: 45 minutes Design: Researchers-driven Training Support: Not reported	PE as usual	Attitudes Self-efficacy Perceived behavioral control Subjective norm Group norm	Results reported that when children who were exposed to threat-framed messages played Wii exergames during PE lessons, they indicated more positive, self-efficacy, and perceived behavioral control than those who underwent regular PE lessons and were exposed to the same message. Among children playing Wii, the threat and coping frames had similar effects on the degree of message influence on physical activity attitudes and beliefs.	Not reported
Mayorga-Vega et al. (2012)	CRCT	PE	*N* = 75 (38 INT; 37 CONT) from 2 intervention classes and 2 control classes in 1 school in Spain Sex F/M (n): 34/41 M_{age}: 11.1	Name: Physical activity and health intervention Total duration: 8 weeks Frequency: 2 times per week Length: 50 minutes Design: Researchers-driven Training Support: Not reported	PE as usual	Global physical self-concept Self-esteem	The results reported overall significant interaction effect on the Physical Self-Description Questionnaire scales ($p<.05$).	Not reported
Mendelson et al. (2010)	RCT	PADS	*N* = 87 (51 INT; 36 CONT) from 2 intervention schools and 2 control schools in an urban area in the USA Sex F/M (n): 59/28 M_{age}: 9.7	Name: Holistic Life Foundation Mindfulness and Yoga Program Total duration: 12 weeks Frequency: 4 times per week Length: 45 minutes Design:	Not reported	Positive affect Negative affect Depression Trust in friends Communication with friends Teacher affiliation Dissatisfaction with teacher Involuntary engagement Rumination Intrusive thoughts Emotional arousal Impulsive action	The INT group reported significant improvements on the overall scale of involuntary engagement when compared to the CONT group ($p<.001$). Significant differences were shown on three of the five subscales of this factor, including rumination ($p<.01$), intrusive thoughts ($p<.05$) and emotional arousal ($p<.01$), as well as a trend in the predicted direction for impulsive action ($p= 0.07$) and physiologic arousal ($p= 0.07$).	Not reported

				Researchers-driven Training Support: Not reported		Physiologic arousal	Although other differences were not statistically significant, adjusted post-intervention means indicated a pattern consistent with predictions for depressive symptoms (7.02 in the INT group versus 7.62 in the CONT group) and negative effect (28.80 in the INT group versus 29.93 in the CONT group). The groups did not differ significantly with respect to changes in positive effect or in relationships with peers and teachers, but there was a trend for CONT group members to report having more trust in friends than INT group members (p=.06).	
Moreno-Murcia & Sánchez-Latorre (2016)	QE	PE	*N* = 145 (94 INT; 54 CONT) from 4 INT groups and 2 CONT groups in 2 schools in Spain Sex F/M (n): 74/71 Age: 10-12	Name: Not reported Total duration: 16 weeks (total 21 sessions) Frequency: 2 times per week Length: Not reported Design: Researchers-driven Training Support: 12-hour training by research team	PE as usual	Perception of autonomy support in class Psychological mediators Motivation Importance and usefulness of physical education Intention to do Physically Active Teacher care	Results indicated that the INT group experienced significant increases in autonomy, intrinsic motivation, importance of PE at pre- and post-assessment.	Observation
Olive et al. (2019)	CRCT	PE	*N* = 821 (445 INT; 376 CONT) from 13 INT schools and 16 CONT schools in a local urban area in Australia Sex F/M (n): 376/445 Age: 7, 8, & 12	Name: Specialist-taught PE intervention Total duration: 4 years Frequency: 2 times per week Length: 50 minutes Design: Researchers-driven Training Support:	PE as usual	Body image: self-perception, others-perceptions, physical attributes, & appearance Depression: negative mood, interpersonal, ineffective, & anhedonia Stress: daily hassles, relationship with parents, transition and change, school	Children of the INT group indicated a 0.71-unit decrease in body dissatisfaction compared to a 3.01-unit increase in CONT group children (p= .042). A mean decrease in depressive symptoms (ineffectiveness), which was 0.27 units more than the CONT group (p= .005). No significant differences in depression and experiences of stress were found between the CONT and INT groups.	Not reported

				By qualified professional organization		problem, & family dissonance/upheaval		
Olive et al. (2020)	QE	PABAS	*N* = 29 (17 INT; 12 CONT) from 1 INT school and 1 CONT school in an urban area in the USA Sex F/M (n): 17/12 Age: 7-12	Name: Physical Activity and Learning Program Total duration: 4 weeks Frequency: 3 times per week Length: 60 minutes Design: Teachers-driven Training support: Not reported	Normal programing as usual	Students' self-perceptions: social and emotional	The INT group's SEL scale variables reported a higher overall score as opposed to the CONT group.	Observation, fidelity checklist, and rubric
PÉRez-OrdÁS et al. (2020)	QE	PE	*N* = 210 (103 INT; 107 CONT) from 4 intervention classes and 4 control classes in 4 schools in Spain Sex F/M (n): 110/100 Age: 10-12	Name: Not reported Total duration: 8 months Frequency: 2 times per week Length: 55 minutes Design: Teacher-driven Training Support: 30-hour training by research team	PE as usual	Physical and verbal aggression Social responsibility behaviors	Significant decreases in physical and verbal aggression's total score (p=.028) and in verbal aggression (p= .003) were shown in the lower-middle socioeconomic context INT group. Verbal aggression was significantly reduced in girls in the INT group (p=.022). Regarding social responsibility, improvements were reported in the INT group.	Bi-weekly meetings
Pesce et al. (2013)	QE	PE	*N* = 125 (63 INT; 62 CONT) from 8 INT classes in 4 schools and 4 CONT classes in 4 schools in an urban area in Italy Sex F/M (n): Not reported Age:	Name: Not reported Total duration: 8 months Frequency: 1 time per week Length: 60 minutes	PE as usual	Task orientation Ego orientation Social self-efficacy Perceived physical ability	The results reported that the multi-sports approach to PE induced more pronounced improvements in aerobic fitness and kinesthetic discrimination ability, as well as small but significant improvements in task orientation, ego orientation and social self-efficacy when compared to the CONT group.	Not reported

			10-11	Design: Researchers-driven Training Support: Not reported				
Shachar et al. (2016)	QE	PABAS	*N* = 330 from 21 schools in underprivileged peripheral geographical areas in Israel Sex F/M (n): 59/271 Grade: 3rd-6th	Name: Sports intervention Total duration: 23 weeks Frequency: 5 times per week Length: 60 minutes Design: Certified sports trainers-driven Training Support: Not reported	No control group	Self-control skills Positive emotion Negative emotion Anger Hostility Physical aggression Teacher-rated child behavior	Results of structural equation modeling reported that self-control skills gains were linked to changes in hostile thoughts, as mediated by changes in both positive and negative emotions. Changes in hostile thoughts were linked to changes in physical aggression through the mediation of changes in anger. Among girls, changes in self-control skills were directly linked to changes in physical aggression (with no indirect effect), whereas among boys, changes in self-control skills were indirectly linked to changes in physical aggression, through changes in positive and negative emotions.	Not reported
Wong et al. (2016)	QE	PABAS	*N* = 877 (524 INT; 353 CONT) from 14 INT community centers and 10 CONT community centers in an urban aera in the USA Sex F/M (n): 240/637 Age: 10.2±1	Name: Healthy Kids-Houston Program Total duration: 18 weeks Frequency: 2 times per week Length: 90 minutes Design: Researchers-driven Training Support: Not reported	After school program as usual	Scholastic competence Social acceptance Athletic competence Physical appearance Behavioral conduct Global self-worth in children	Significant improvements ($p \leq .02$) were indicated in the amount of exercise that a child perceived to be required to offset a large meal and in several key self-esteem scores.	Not reported

Notes: CI: Confidence Interval; CONT=Control; CRCT=Cluster Randomized Controlled Trial; ES: Effect Size; FCE=Family and Community Engagement; FMS=Fundamental Motor Skill; INT=Intervention; PABAS=Physical Activity Before and After School; PADS=Physical Activity During School; PE=Physical Education; QE=Quasi Experimental; RCT= Randomized Controlled Trial; SEL=Social and Emotional Learning; SSIS□RS= Social Skills Improvement System Rating Scale; TPSR=Teaching Personal and Social Responsibility.