Preventive Medicine 155 (2022) 106948

Contents lists available at ScienceDirect

Preventive Medicine

journal homepage: www.elsevier.com/locate/ypmed

Review Article

The effects of motor skill and physical activity interventions on preschoolers' cognitive and academic skills: A systematic review

Pinja Jylänki [a,*], Theo Mbay [a], Airi Hakkarainen [b], Arja Sääkslahti [c], Pirjo Aunio [a]

[a] Faculty of Educational Sciences, University of Helsinki, Siltavuorenpenger 5, 00170 Helsinki, Finland
[b] Open University, University of Helsinki, Yliopistonkatu 3, 00100 Helsinki, Finland
[c] Faculty of Sport and Health Sciences, University of Jyväskylä, Keskussairaalantie 4, 40600 Jyväskylä, Finland

ARTICLE INFO

Keywords:
Motor skills
Physical activity
Cognition
Academic skills
Early intervention
Systematic review

ABSTRACT

The present systematic review aimed to investigate the methodological quality and the effects of fundamental motor skills and physical activity interventions on cognitive and academic skills in typically developing 3 to 7-year-old children. The review was conducted and reported in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement. A literature search was carried out in April 2020 using seven electronic databases. The methodological quality of the studies was assessed with the Effective Public Health Practice Project (EPHPP) Quality Assessment Tool. Cohen's *d* effect size calculations and post hoc power analyses were conducted for the included studies. A total of 35 studies, representing 2472 children met the inclusion criteria. Two of the studies demonstrated a strong methodological quality, while 24 were considered as methodologically weak. The majority (71%) of the included studies demonstrated the beneficial effects of the intervention on cognitive and academic skills. The most evidence was found for executive functions, language, and numeracy, and the effects were largest in enhancing memory. The effects were larger on cognitive and academic skills in the combined interventions compared to only fundamental motor skill and physical activity interventions, while fundamental motor skill interventions had larger effects than physical activity interventions. These findings indicate that it may be possible to support typically developing preschoolers' cognitive and academic learning with fundamental motor skill and physical activity interventions. However, most of the studies in this field have a weak methodological quality and thus, the presented evidence was considered weak in nature.

1. Introduction

Early childhood is a critical phase in the development of fundamental motor (FMS; Gallahue and Ozmun, 2002), cognitive (Burger, 2010), and academic skills (e.g., early numeracy and language skills; Duncan et al., 2007); which has shown to predict later school adjustment (Bart et al., 2007) as well as academic achievements (Duncan et al., 2007). FMS can be seen as the "building blocks" of more complex movement skills that are required in games or other physical activities (Logan et al., 2018), and consists of balance (e.g. rolling, turning), manipulative (e.g., catching, kicking), and locomotor movement skills (e.g., running, jumping; Gallahue and Ozmun, 2002). Physical activity (PA) includes all voluntary bodily movements produced by skeletal muscles that result in energy expenditure (Caspersen et al., 1985). When young children perform FMS, which are produced by large muscles of the body, the energy expenditure and PA is high (Malina et al., 2004; Gallahue and Ozmun, 2002). Thus, among young children there is very close relationship between FMS and PA (Stodden et al., 2008; Jones et al., 2020). In recent years, the research has increasingly focused on the effects of FMS and PA on cognitive and academic skills and has shown a positive relationship between these skills in preschoolers (Planinšec, 2002). A potential explanation for the close association between FMS and cognitive skills is the co-activation of the same brain areas (i.e., cerebellum and prefrontal cortex; Diamond, 2000). It has also been suggested that the effects of PA on cognitive and academic skills may be mediated by other factors such as physical fitness, health, and psycho-social factors (Tomporowski et al., 2011). Indeed it has been shown that physical fitness and cognitive skills are related in young children (Ruiz-Hermosa et al., 2020), however the relationship between PA and physical fitness will strengthen in the later childhood (Stodden et al., 2008).

Although the positive associations between FMS and PA with various

* Corresponding author.
E-mail address: pinja.jylanki@helsinki.fi (P. Jylänki).

https://doi.org/10.1016/j.ypmed.2021.106948
Received 24 June 2021; Received in revised form 18 November 2021; Accepted 27 December 2021
Available online 30 December 2021

health indicators (i.e., cardiometabolic health, bone and skeletal health, psychosocial health, self-esteem, and cognitive development) have been demonstrated (Carson et al., 2017; Rasmussen and Laumann, 2013), children have become less physically active during recent years and almost half of preschool-aged children do not meet the recommended daily PA guidelines (i.e., at least 60 min of physical activity and several hours of unstructured play per day; National Association for Sport and Physical Education, 2002; Tucker, 2008). As a consequence, children's FMS, especially coordinative skills, have been declining (Roth et al., 2010), and children spent excessive time being sedentary (e.g., engage in screen-based entertainments; Hinkley et al., 2012). This increased sedentary time during early childhood has been negatively associated with children's cognitive and academic skills (Carson et al., 2015). Due to the widespread attendance, preschools exhibit a crucial role as a supporter of beneficial development in FMS, PA habits (Tonge et al., 2016) and cognitive skills (Burger, 2010).

Previous systematic reviews that have examined the effects of FMS or PA interventions on cognitive and academic skills, have concentrated mainly on school-aged children (i.e., 5–12-year-olds) and included only PA interventions (Watson et al., 2017), have used inadequate quality assessment and included a small number of studies (Zeng et al., 2017) or have not conducted effect size calculations (Carson et al., 2016). Thus, the present systematic review aimed to investigate 1) the methodical quality and 2) the effects of FMS and PA interventions on cognitive and academic skills in 3 to 7-year-old children. In addition, the effectiveness of only FMS and/or PA interventions were compared with combined interventions (FMS and/or PA combined with cognitive or academic skill practices, e.g., FMS practices combined with story reading; Bedard et al., 2017). Because this systematic review concentrated in 3 to 7-year-old children, both FMS and PA interventions were included. In order to separate FMS and PA interventions from each other, in this review PA interventions were conceptualized as interventions that do not include any specific skill practices but rather focuses on increasing PA.

2. Methods

The systematic review was conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement, which can be used especially in evaluations of intervention studies (Moher et al., 2009). The following steps were carried out in this systematic review:

Step 1: A literature search including abstract rating and full-text screening based on the predetermined eligibility criteria.

Step 2: Methodological quality of the studies was assessed with the Effective Public Health Practice Project Quality Assessment Tool for Quantitative Studies (EPHPP).

Step 3: Cohen's d effect sizes were calculated and post hoc power analyses were conducted to determine the statistical power of the selected studies.

2.1. Eligibility criteria

Peer-reviewed intervention studies, investigating the effects of FMS or PA interventions on preschoolers' cognitive or academic skills, that were published in English were eligible to be included in the present systematic review. Eligibility criteria included the following:

The main content of studies: intervention studies with at least one program including only FMS and/or PA practices or a program that combined FMS and/or PA with cognitive skill practices.

Target population: typically developing children between the ages 3 and 7. Based on the participant description, children with special needs, low socioeconomic status, or migrant background were excluded from the present review since the lower baseline scores might have affected the generalizability of the results.

Study design: all intervention designs, except case studies, were included; i.e., randomized controlled trials (RCT), controlled clinical trials (CCT), pre-post designs with a control group (PPD), and acute interventions.

Outcome measures: intervention effects were assessed with cognitive or academic skills (i.e., executive functions, cognitive performance, language skills, numeracy, or memory) measures.

Type of publication: published peer-reviewed journal articles written in English. All publications before April 16th, 2020 were searched.

The systematic literature search was carried out in April 2020 by two authors using seven electronic databases including ERIC, Scopus, Web of Science, PsycINFO, CINAHL, PubMed and SPORTDiscus. Search terms were designed in accordance with the PICO framework (Melnyk and Fineout-Overholt, 2005). The final search string consisted of the following: "early education" OR child* AND motor* OR "physical activity" AND intervention OR program* OR treatment OR training OR instruction AND cognit* OR academic*. When possible, additional search filters were used to exclude studies that investigated children older than 7 years.

2.2. Study selection

Two authors performed the article selection according to the predetermined eligibility criteria. Initially, the articles were screened based on the title and the abstract. In terms of inclusion, the abstracts were coded as "yes", "maybe" or "no". The inter-rater agreement was determined during the abstract rating processes by calculating Cohen's weighted kappa. Both authors rated the first 40% (n = 2266) of the abstracts independently, after which the inter-rater agreement was 0.718, which can be considered as a good agreement between the raters (Byrt, 1996). After the abstract screening, all of the eligible articles underwent full-text screening. Both authors decided independently whether to "exclude", "include" or "maybe" include each article. Disagreements between the two authors were solved in consensus meetings with all of the five authors.

2.3. Methodological quality

The methodological quality of the eligible studies was assessed with the EPHPP tool (National Collaborating Centre for Methods and Tools, 2008; Thomas et al., 2004). The tool is suitable for evaluating the quality of a variety of study designs (e.g., RCTs and PPDs; Armijo-Olivo et al., 2012) and it has been used previously in systematic reviews in this particular field (e.g., Norris et al., 2015). The inter-rater agreement has shown to be more consistent in the EPHPP tool comparison to the Cochrane Collaboration Risk of Bias Tool (Armijo-Olivo et al., 2012). Two authors rated each study procedure as "strong", "moderate" or "weak". Final ratings were formed based on six sections (selection bias, study design, confounders, blinding, data collection methods, and withdrawals and drop-outs) with the following criteria: studies with no weak ratings and at least four strong ratings were considered as "strong"; studies with less than four strong ratings and one weak rating were considered as "moderate"; studies with two or more weak ratings were considered as "weak". For more detailed criteria see Quality Assessment Tool for Quantitative Studies Dictionary (National Collaborating Centre for Methods and Tools, 2008). Disagreements were solved in consensus meetings with all of the authors.

2.4. Data extraction

Data were extracted from the eligible studies independently by two authors. Extracted data included the geographical location, study design, sample size, children's age and gender, cognitive and academic skill measures, intervention exposure, intervention details (only FMS and/or PA interventions and combined interventions), control conditions, and data for effect size calculations. If missing data was encountered, the corresponding author was contacted in order to receive the required information.

2.5. Effect size calculations

Cohen's *d* effect sizes (Cohen, 1988) were calculated to allow for the quantification and comparison of the effects across the studies. Effect sizes were calculated for the studies that demonstrated significant effects and provided sufficient information (i.e., pre and post-scores as well as associated standard deviations or standard errors). If a study demonstrated significant effects for multiple outcomes, all of them were included. Cohen's *d* effect sizes of <0.2, 0.2, 0.5, and 0.8, correspond to trivial, small, medium, and large effects, respectively (Cohen, 1988).

Between-group effects were calculated followingly.

$$ES_{(d)} = \frac{(M_{post,E} - M_{pre,E}) - (M_{post,C} - M_{pre,C})}{SDpooled_{pre}}$$

where

$$SDpooled_{pre} = \sqrt{\frac{((n_E - 1)\ SD_{pre2,E} + (n_C - 1)\ SD_{pre2,C})}{(n_E + n_C - 2)}}$$

, and within-group effects were calculated as

$$ES_{(d)} = \frac{(M_{post} - M_{pre})}{SD_{pre}}$$

For the studies assessing the learning of previously unknown items, i.e., where the baseline score was approximating zero (Padial-Ruz et al., 2019; Mavilidi et al., 2015; Mavilidi et al., 2016a; Mavilidi et al., 2017a; Mavilidi et al., 2018a; Toumpaniari et al., 2015), the effect sizes were calculated as:

$$ES_{(d)} = \frac{(M_{post,E} - M_{post,C})}{SDpooled_{post}}$$

where

$$SDpooled_{post} = \sqrt{\frac{((n_E - 1)\ SD_{post2,E} + (n_C - 1)\ SD_{post2,C})}{(n_E + n_C - 2)}}$$

ES (d) = Cohen's d effect size
Mpost = mean post score
Mpre = mean pre score
E = experimental group
C = control group
SDpooled = pooled standard deviation
n = sample size.

2.6. Power analyses

Power calculations were carried out with G*power 3.1.9.6 (Faul et al., 2007). If a study had multiple groups, the power calculations were conducted on a sub-group basis in order to determine the power of specific group comparisons. Type 1 error probability (α) was computed as 0.05, corresponding to a significance level of 5%. A medium effect size (0.5) was used as the reference point to establish observed power for each outcome and type 2 error probability (β) of 0.2, corresponding to a power of 0.8 (1 – β), or 80%, was selected as the cut-off point for adequate power (Schulz and Grimes, 2005).

3. Results

3.1. Search results

In the systematic literature search a total number of 10,995 articles, 5665 after duplicates were removed, were found. Additional 4 articles were found through other sources. A total number of 5669 articles were screened based on the abstract and title. After that, 5471 articles were excluded due to not meeting the eligibility criteria and 198 full-text articles were further assessed for eligibility. Finally, 35 articles were included in the main analysis (see Fig. 1).

3.2. Study characteristics and population

All study characteristics, including authors, study design, participants, outcome measures, intervention duration, and intervention/control activities, are presented in detail in Supplementary Table S1. Included studies were published between March 2010 (Ellemberg and St-Louis-Deschênes, 2010) and February 2020 (Botha and Africa, 2020). The studies were conducted in 17 countries, representing North America, Europe, Asia, Australia, and Africa (see Supplementary Table S1). Participants' ages ranged from 3.3 years (Bedard et al., 2020) to 7.8 years (Ellemberg and St-Louis-Deschênes, 2010). One study (Ellemberg and St-Louis-Deschênes, 2010) included only boys and other studies included both girls and boys (i.e., 37.8% (Jarraya et al., 2019) - 78.6% (Gao et al., 2019) of the participants were boys). The total sample in the 35 studies included 2472 children.

3.3. Intervention characteristics

In total, 35 articles with 46 intervention programs were included in the present review. Eleven studies (Barnard et al., 2014; Callcott et al., 2015a; Chang et al., 2013; Duncan et al., 2019; Jarraya et al., 2019; Mavilidi et al., 2015; Mavilidi et al., 2016a; Mavilidi et al., 2017a; Mavilidi et al., 2018a; Shoval et al., 2018; Xiong et al., 2019a) included two programs that met the eligibility criteria. Seventeen of the intervention programs focused on FMS practises, 15 programs on PA practises and 14 programs on both FMS and PA practises. Twenty-four of the interventions were only FMS and/or PA programs and 22 interventions combined FMS and/or PA practices with cognitive and/or academic skill practices in the program. Intervention duration varied from one week (Mavilidi et al., 2016a; Mavilidi et al., 2016b) to one academic year (Bala et al., 2013; Have et al., 2018). Two of the interventions (Ellemberg and St-Louis-Deschênes, 2010; Stein et al., 2017) were acute interventions, and in three studies (Callcott et al., 2015a; Habibi et al., 2014; Raney et al., 2017; Shoval et al., 2018) the intervention duration was not reported. The duration of one intervention session varied from five minutes (St. Laurent et al., 2018) to three hours (Shoval et al., 2018), and sessions were held once a week (Bedard et al., 2017; Bedard et al., 2020; Fisher et al., 2011; Mavilidi et al., 2017a; Mavilidi et al., 2017b; Mavilidi et al., 2018a; Mavilidi et al., 2018b; Pienaar et al., 2011) to six times a week (Have et al., 2018). Outcome measures were assigned to six categories based on the provided description of the measure. The categories were as follows: executive functions (e.g., Eriksen Flanker task (Beck et al., 2016; Chang et al., 2013; Have et al., 2018), visuo-spatial short-term memory; Have et al., 2018), language (e.g., the ESSI reading and spelling tests; Bala et al., 2013; Shoval et al., 2018), cognitive skill (e.g., Raven Progressive Matrices assessment test; Bala et al., 2013; Shoval et al., 2018), numeracy (e.g., Mathematics Achievement Test; Shoval et al., 2018), memory (e.g., free-recall and cued recall tests; Mavilidi et al., 2015; Mavilidi et al., 2017a) and other tests (curriculum test; Raney et al., 2017).

3.4. Methodological quality

The methodological quality was determined based on the following factors: selection bias, study design, confounders, blinding, data collection methods, and withdrawals and drop-outs. Only two of the included 35 studies (6%) demonstrated strong methodological quality, while nine of the studies (26%) had moderate quality and 24 of the studies (69%) were considered as methodologically weak. The rating for each section as well as the overall quality of the studies are presented in Table 1.

Of the 35 studies, 30 were CCTs (i.e., quasi-experimental designs with a control group and RCTs that did not report the randomization process) and five studies were RCTs. In 19 studies (Bala et al., 2013; Barnard et al., 2014; Battaglia et al., 2019; Beck et al., 2016; Botha and Africa, 2020; Callcott et al., 2015a; Duncan et al., 2019; Fisher et al., 2011; Have et al., 2018; Mavilidi et al., 2015; Mavilidi et al., 2016a; Mavilidi et al., 2017a; Mavilidi et al., 2018a; Sánchez-López et al., 2019; Shoval et al., 2018; Stein et al., 2017; Toumpaniari et al., 2015; Xiong

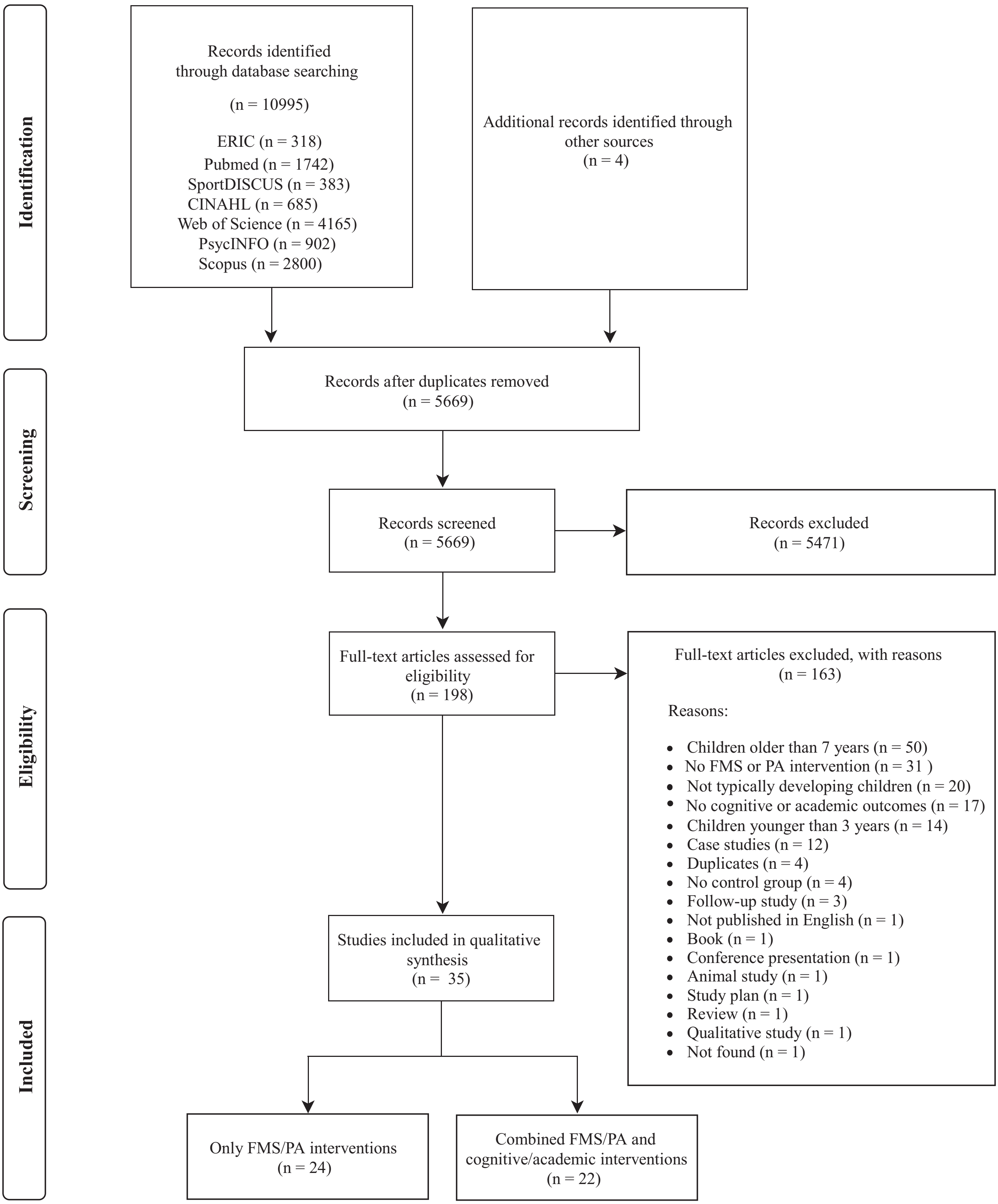

Fig. 1. PRISMA flow diagram of the stages associated with the systematic selection of studies.

Table 1
Methodological quality of the included studies.

Reference	Selection bias	Study design	Confounders	Blinding	Data collection methods	Withdrawal and dop outs	Overall quality scores
Bala et al., 2013	Moderate	Strong	Strong	Moderate	Weak	Weak	Weak
Barnard et al., 2014	Moderate	Strong	Weak	Moderate	Weak	Weak	Weak
Battaglia et al., 2019	Moderate	Strong	Strong	Weak	Weak	Weak	Weak
Beck et al., 2016	Moderate	Strong	Strong	Moderate	Weak	Strong	Moderate
Bedard et al., 2017	Weak	Strong	Strong	Moderate	Moderate	Strong	Moderate
Bedard et al., 2020	Weak	Strong	Strong	Moderate	Moderate	Strong	Moderate
Botha and Africa, 2020	Moderate	Strong	Strong	Strong	Weak	Strong	Moderate
Callcott et al., 2015a	Moderate	Strong	Weak	Moderate	Strong	Weak	Weak
Chang et al., 2013	Weak	Strong	Strong	Moderate	Weak	Weak	Weak
Duncan et al., 2019	Moderate	Strong	Weak	Moderate	Moderate	Weak	Weak
Ellemberg and St-Louis-Deschênes, 2010	Weak	Strong	Strong	Moderate	Weak	Strong	Weak
Fischer et al., 2011	Weak	Strong	Weak	Moderate	Weak	Strong	Weak
Fisher et al., 2011	Weak	Strong	Strong	Strong	Weak	Strong	Weak
Gao et al., 2019	Weak	Strong	Strong	Strong	Strong	Strong	Moderate
Habibi et al., 2014	Weak	Strong	Strong	Moderate	Weak	Strong	Weak
Have et al., 2018	Moderate	Strong	Strong	Strong	Moderate	Strong	Strong
Jarraya et al., 2019	Weak	Strong	Strong	Strong	Weak	Weak	Weak
Lai et al., 2020	Weak	Strong	Weak	Moderate	Weak	Weak	Weak
Lee et al., 2020	Weak	Strong	Strong	Moderate	Strong	Strong	Moderate
Mavilidi et al., 2015	Moderate	Strong	Weak	Moderate	Weak	Strong	Weak
Mavilidi et al., 2016a	Moderate	Strong	Weak	Moderate	Weak	Strong	Weak
Mavilidi et al., 2017a	Moderate	Strong	Strong	Moderate	Weak	Strong	Moderate
Mavilidi et al., 2018a	Moderate	Strong	Weak	Moderate	Weak	Strong	Weak
Padial-Ruz et al., 2019	Weak	Strong	Weak	Moderate	Weak	Strong	Weak
Pienaar et al., 2011	Weak	Strong	Weak	Moderate	Strong	Weak	Weak
Ramah, 2014	Weak	Strong	Strong	Moderate	Weak	Weak	Weak
Raney et al., 2017	Weak	Strong	Weak	Moderate	Weak	Weak	Weak
Sánchez-López et al., 2019	Moderate	Strong	Strong	Moderate	Moderate	Moderate	Moderate
Shoval et al., 2018	Moderate	Strong	Strong	Moderate	Moderate	Weak	Moderate
Stein et al., 2017	Moderate	Strong	Strong	Moderate	Weak	Weak	Moderate
St. Laurent et al., 2018	Weak	Strong	Strong	Moderate	Weak	Moderate	Weak
Toumpaniari et al., 2015	Moderate	Strong	Weak	Moderate	Weak	Weak	Weak
Wen et al., 2018	Weak	Strong	Strong	Strong	Strong	Weak	Weak
Xiong et al., 2017	Moderate	Strong	Weak	Moderate	Weak	Strong	Weak
Xiong et al., 2019a	Moderate	Strong	Strong	Moderate	Strong	Strong	Strong

Note. Some modifications were made to the EPHPP tool to solve misunderstandings between the raters.
Study design: Studies that used quasi-experimental design were coded as CCT.
Confounders: The confounders of interest included age, gender, health status and pre-intervention score.
Blinding: In question 2 "Were the study participants aware of the research question?" we chose to code "no", if there was no mention that participants were aware of the research question. This decision was made based on the young age of the participants.
Data collection methods: The outcome of interest (cognitive or academic measurement) was evaluated. Methods were coded to be "valid" if the validity was mentioned in the article or if there was a citation to test manual or other article where the validity was reported. Some well known methods were seen as valid methods without mentioning (e.g., Wechsler Intelligence Scale for Children or Bayley Scales of Infant and Toddler Development). Methods were coded as "reliable" only, if the reliability was measured and reported in that specific data.
Withdrawals and drop-outs: In question 1, "Were withdrawals and drop-outs reported in terms of numbers and/or reasons per group?", if both numbers and reasons are reported then it was code as "yes", otherwise "no". Withdrawals and drop-outs were considered as children that did not finish the intervention, i.e. not missing data.

et al., 2017; Xiong et al., 2019a), the participants were referred from a source (e.g., preschool) in a systematic manner, and thus the participants were considered only somewhat likely to be representative of the target population. None of the studies referred the participants through randomization. In 18 studies (Bedard et al., 2017; Mavilidi et al., 2017a; Chang et al., 2013; Have et al., 2018; Botha and Africa, 2020; Bala et al., 2013; Bedard et al., 2020; Ellemberg and St-Louis-Deschênes, 2010; Ramah, 2014; Jarraya et al., 2019; Habibi et al., 2014; Battaglia et al., 2019; Gao et al., 2019; Lee et al., 2020; Wen et al., 2018; Duncan et al., 2019; Xiong et al., 2019a; Stein et al., 2017) there were not any important confounders between groups (i.e., participants' age, gender, health status or pre-intervention score) and in four studies (Beck et al., 2016; Fisher et al., 2011; Shoval et al., 2018; St. Laurent et al., 2018) the confounders were 80–100% controlled. Only six studies (Botha and Africa, 2020; Fisher et al., 2011; Gao et al., 2019; Have et al., 2018; Jarraya et al., 2019; Wen et al., 2018) reported that the outcome assessors were not aware, and two studies (Battaglia et al., 2019; Raney et al., 2017) reported that the outcome assessors were aware of the intervention exposure. The validation of the data collection methods was reported in 12 studies (Bedard et al., 2017; Bedard et al., 2020; Callcott et al., 2015a; Duncan et al., 2019; Gao et al., 2019; Have et al., 2018; Lee et al., 2020; Pienaar et al., 2011; Sánchez-López et al., 2019; Shoval et al., 2018; Wen et al., 2018; Xiong et al., 2019a) and of these, only six studies (Callcott et al., 2015a; Gao et al., 2019; Lee et al., 2020; Pienaar et al., 2011; Wen et al., 2018; Xiong et al., 2019a) reported that the data collection methods demonstrated good reliability in that particular data. Most of the studies (Beck et al., 2016; Bedard et al., 2017;: Bedard et al., 2020; Botha and Africa, 2020; Ellemberg and St-Louis-Deschênes, 2010; Fischer et al., 2011; Fisher et al., 2011; Gao et al., 2019; Have et al., 2018; Habibi et al., 2014; Lee et al., 2020; Mavilidi et al., 2015; Mavilidi et al., 2016a; Mavilidi et al., 2017a; Mavilidi et al., 2018a; Padial-Ruz et al., 2019; Sánchez-López et al., 2019; St. Laurent et al., 2018; Xiong et al., 2017) reported both withdrawals and drop-outs in terms of numbers and reasons. Thirty-one (89%) of the included studies were found to be underpowered to detect a medium effect size, and only nine of the included studies reported having conducted a priori power analyses (Raney et al., 2017; Jarraya et al., 2019; Habibi et al., 2014; Gao et al., 2019; Lai et al., 2020; Stein et al., 2017; Wen et al., 2018; Xiong et al., 2019a; Have et al., 2018); of which only one was adequately powered (Have et al., 2018) and five stated low statistical power as a limitation of the study (Gao et al., 2019; Jarraya et al., 2019; Lai et al., 2020; Wen et al., 2018; Xiong et al., 2019a).

3.5. Effect sizes

In terms of the intervention effects (see Supplementary Table S2), twelve of the included studies assessed executive functions as an outcome variable; three with combined interventions (Gao et al., 2019; Have et al., 2018; Beck et al., 2016) and nine with FMS and/or PA only interventions (Chang et al., 2013; Ellemberg and St-Louis-Deschênes, 2010; Fisher et al., 2011; Jarraya et al., 2019; Lai et al., 2020; Stein et al., 2017; Wen et al., 2018; Xiong et al., 2017; Xiong et al., 2019a).[1] Two of the three (67%) combined interventions (Gao et al., 2019; Beck et al., 2016), and eight out of nine (89%) FMS and/or PA only interventions (Chang et al., 2013; Ellemberg and St-Louis-Deschênes, 2010; Fisher et al., 2011; Jarraya et al., 2019; Lai et al., 2020; Stein et al., 2017; Xiong et al., 2017; Xiong et al., 2019a) demonstrated significant effects. The effect sizes were medium ($d = 0.50$–0.68, $\bar{x}$ 0.59) for FMS and/or PA only interventions and ranged from trivial to large for combined interventions ($d = 0.12$–1.47, $\bar{x}$ 0.68). Out of the two null-findings (Wen et al., 2018; Have et al., 2018) one study (50%; Wen et al., 2018) was underpowered to detect a medium effect.

A total of ten studies assessed language skills as an outcome variable; four were FMS and/or PA only interventions (Barnard et al., 2014; Battaglia et al., 2019; St. Laurent et al., 2018; Habibi et al., 2014), three were combined interventions (Bedard et al., 2017; Bedard et al., 2020; Botha and Africa, 2020), and three included both (Callcott et al., 2015a; Duncan et al., 2019; Shoval et al., 2018). Three (Duncan et al., 2019; Barnard et al., 2014; Callcott et al., 2015a) of the seven (43%) FMS and/or PA only interventions and four (Botha and Africa, 2020; Callcott et al., 2015a; Duncan et al., 2019; Shoval et al., 2018) out of six (67%) combined interventions demonstrated significant effects. The effect sizes ranged from trivial to small ($d = 0.16$–0.25, $\bar{x}$ 0.20) for FMS and/or PA only interventions and from small to large for combined interventions ($d = 0.31$–2.2, $\bar{x}$ 0.90). All of the null findings (Battaglia et al., 2019; St. Laurent et al., 2018; Shoval et al., 2018; Habibi et al., 2014; Bedard et al., 2017) were underpowered to detect a medium effect. Furthermore, three of the studies (Callcott et al., 2015a; Duncan et al., 2019; Shoval et al., 2018) compared FMS and/or PA only interventions to combined interventions and demonstrated small effects ($d = 0.23$–0.36, $\bar{x}$ 0.31) in favor of the combined interventions.

Seven studies included a numeracy measure as an outcome variable; one was a FMS and/or PA only intervention (Barnard et al., 2014), five were combined interventions (Mavilidi et al., 2018a; St. Laurent et al., 2018; Have et al., 2018; Beck et al., 2016; Fischer et al., 2011) and one included both (Shoval et al., 2018). None of the FMS and/or PA only interventions, and five (83%) of the six combined interventions (Mavilidi et al., 2018a; Shoval et al., 2018; Have et al., 2018; Beck et al., 2016; Fischer et al., 2011) demonstrated significant effects, which ranged from small to large ($d = 0.38$–0.94, $\bar{x}$ 0.57). Furthermore, one of the studies compared FMS and/or PA only to a combined intervention (Shoval et al., 2018) and found a large effect ($d = 0.89$) in favor of the combined intervention. Of the three null findings (Barnard et al., 2014; St. Laurent et al., 2018; Shoval et al., 2018), one was underpowered to detect a medium effect (St. Laurent et al., 2018).

Five studies (Padial-Ruz et al., 2019; Mavilidi et al., 2017a; Mavilidi et al., 2016a; Mavilidi et al., 2015; Toumpaniari et al., 2015) investigated the effects of adding FMS and/or PA on a learning task to enhance memory. All of the studies demonstrated large significant effects ($d = 1.23$–2.2, $\bar{x}$ 1.52) when the FMS and/or PA was directly related to the learning task. Three of these studies also included a group performing FMS and/or PA unrelated to the learning activity (Mavilidi et al., 2015; Mavilidi et al., 2016a; Mavilidi et al., 2017a), and two (67%) found significant large effects ($d = 0.82$–1.20, $\bar{x}$ 1.0; Mavilidi et al., 2016a; Mavilidi et al., 2017a) while one did not find any significant effects (Mavilidi et al., 2015). The one null finding (Mavilidi et al., 2015) was underpowered to detect a medium effect.

A total of seven studies assessed cognitive skills as an outcome variable, of which six were FMS and/or PA only interventions (Habibi et al., 2014; Bala et al., 2013; Pienaar et al., 2011; Lee et al., 2020; Ramah, 2014; Sánchez-López et al., 2019) and one included a FMS and/or PA only intervention as well as a combined intervention (Shoval et al., 2018). Of the seven FMS and/or PA only interventions, three (43%) demonstrated significant effects (Sánchez-López et al., 2019; Shoval et al., 2018; Ramah, 2014), with effect sizes ranging from small to large ($d = 0.40$–1.3, $\bar{x}$ 0.77). The one combined intervention found a significant medium effect ($d = 0.60$) and when compared to FMS and/or PA only, demonstrated a small effect in favor of the combined intervention ($d = 0.35$). All of the four null findings (Bala et al., 2013; Habibi et al., 2014; Pienaar et al., 2011; Lee et al., 2020) were underpowered to detect a medium effect.

Finally, one study used a biology curriculum test as a measure and did not observe a significant impact of a combined intervention (Raney et al., 2017). The one null finding (Raney et al., 2017) was underpowered to detect a medium effect.

3.6. Methodological quality and effect sizes

Methodological quality and effect sizes are presented in Table 2. Methodologically strong studies found medium (Xiong et al., 2019a) and trivial effects (Xiong et al., 2019b) on executive functions and small effects (Have et al., 2018) on numeracy. Large effects were found in 13 studies and seven of these (54%) had a weak, while six (46%) had a moderate methodological quality. Of these, only two studies (Sánchez-López et al., 2019; Shoval et al., 2018) used outcome measures that were shown to be valid, while three studies (Mavilidi et al., 2016a; Mavilidi et al., 2017a; Mavilidi et al., 2018a) used outcome measures that were shown to be reliable. Six of these studies (Jarraya et al., 2019; Stein et al., 2017; Lai et al., 2020; Botha and Africa, 2020; Mavilidi et al., 2015; Toumpaniari et al., 2015) used outcome measures that were neither shown to be valid or reliable. The studies that received a strong rating in the data collection methods section demonstrated medium effects in three studies (Callcott et al., 2015b; Gao et al., 2019; Xiong et al., 2019a), small effects in one study (Callcott et al., 2015b), and trivial effects in one study (Xiong et al., 2019b).

4. Discussion

The present systematic review aimed to investigate the methodological quality and the effects of FMS and PA interventions on cognitive and academic skills in typically developing 3–7-year-old children. These findings demonstrated that only 6% of the included studies had a strong methodological quality, while 69% were considered methodologically weak, which in general was the result of inadequate reporting practices. The majority (71%) of the included studies (i.e., both FMS and/or PA only and combined interventions) demonstrated beneficial effects on cognitive and academic performance of preschoolers. The most evidence was found for the benefits on executive functions, language and numeracy, and the effect sizes were largest in enhancing memory. However, the studies that found large effects had low-to-moderate methodological quality and none of these studies used both validated and reliable outcome measures. In general, combined interventions resulted in beneficial outcomes more consistently than FMS and/or PA only interventions, as well as demonstrated greater effects in magnitude. In the comparison of FMS or PA only interventions, FMS only interventions appeared to have larger effects on cognitive and academic skills than PA only interventions.

[1] Chang et al., 2013 and Xiong et al., 2019a included active control groups and thus, the effects were analysed using a within group analysis.

Table 2
Summary of the relationship between methodological quality and intervention effects.

Intervention	Outcome	Effect size (*d*)				
		Not significant	Trivial	Small	Medium	Large
FMS	Executive functions	Stein et al. 2017 [c]			Stein et al. 2017 [b]	*Jarraya et al. 2019a [c];* Stein et al. 2017 [b]
	Language skills	*Battaglia et al. 2019[c]; Callcott et al. 2015a [c]*		*Duncan et al. 2019a [b]*		
	Cognitive skills	*Bala et al. 2013 [c]; Pienaar et al. 2011 [c];* Lee et al. 2020 [c]		*Ramah et al. 2014 [c]*		
PA	Executive functions	*Wen et al. 2018 [c]*	*Ellemberg et al. 2010 [c]*	*Ellemberg et al. 2010 [c]*		
	Language skills	*Habibi et al. 2014 [c]*				
	Cognitive skills	*Habibi et al. 2014 [c]*				
FMS & PA	Executive functions	*Fisher et al. 2011 [c]; Jarraya et al. 2019b[c]; Lai et al. 2020 [c]*	**Xiong et al. 2019b [c]**	*Chang et al. 2013b [b]*	*Chang et al. 2013a [b]; Fisher et al. 2011 [c]; Xiong et. al. 2017 [c];* **Xiong et al. 2019a [c]**	*Lai et al. 2020 [b]*
	Language skills	*Barnard et al. 2014[a];* Shoval et al. 2018a [c]	*Barnard et al. 2014 [a]*			
	Cognitive skills	Shoval et al. 2018a [c e]			Shoval et al. 2018a [c]	Sánchez-López et al. 2019 [c]
	Numeracy	*Barnard et al. 2014[a];* Shoval et al. 2018a [c]				
FMS combined	Executive functions	Beck et al. 2016 [c]		Beck et al. 2016 [b]	Beck et al. 2016 [b]	
	Language skills	Bedard et al. 2017 [c]; Bedard et al. 2020 [c]; *Callcott et al. 2015b [f]*		*Callcott et al. 2015b [c d] Duncan et al. 2019b [d f]*	Botha et al. 2020 [c]; *Callcott et al. 2015b[c]; Duncan et al. 2019b [b]*	Botha et al. 2020 [c]
	Numeracy	Beck et al. 2016 [c]; *Fischer et al. 2011 [c]*		Beck et al. 2016 [b]	*(Fischer et al. 2011 [c])*	
PA combined	Executive functions	**Have et al. 2018 [c]**				
	Numeracy	*Mavilidi et al. 2018b[c]*		**(Have et al. 2018 [c])**		*Mavilidi et al. 2018a [c]*
	Memory performance	*Mavilidi et al. 2015b[c]*				*Mavilidi et al. 2015a [c]; Mavilidi et al. 2016a [c]; Mavilidi et al. 2016b[c];* Mavilidi et al. 2017a [c]; Mavilidi et al. 2017b [c]; *Toumpaniari et al. 2015[c]*
	Biology curriculum test	*Raney et al. 2017 [c]*				
FMS & PA combined	Executive functions				Gao et al. 2019 [c]	
	Language skills	*St. Laurent et al. 2018 [c]*		Shoval et al. 2018b [c d]		
	Cognitive skills			Shoval et al. 2018b [c d]	Shoval et al. 2018b [c]	
	Numeracy	*St. Laurent et al. 2018 [c]*		Shoval et al. 2018b [c d]		Shoval et al. 2018b [d]

Note. a and b in the end of the citation refers to different intervention programs within one study.
[a] Barnard et al. 2014a and Barnard et al. 2014b analyzed together
[b] Within group analysis
[c] Intervention compared to control
[d] Combined intervention compared to only FMS/PA intevention
[e] Only FMS/PA intevention compared to combined intervention
[f] Combined intervention compared to academic only
Methodological quality based on the EPHPP: *weak*, moderate, **strong**
If it was not possible to calculate effect sizes but the study reported a self-calculated Cohen's d effect size, it was presented in parentheses.

In terms of the intervention effects, the largest effects of FMS and/or PA interventions were found for memory, and these benefits were consistently shown only when the PA was directly related to the learning task (e.g., moving in the manner of a particular animal when learning the names of animals in foreign language), and less consistently when the PA was unrelated to the activity (e.g., running in circles after the learning task). Nonetheless, more research is required to elucidate whether these effects translate to general outcomes. The most commonly assessed outcomes in the interventions were executive functions and language skills. For executive functions, 83% of the interventions demonstrated significant benefits with medium effect sizes, with no difference in terms of the magnitude of the effect in FMS and/or PA only and combined interventions. For language skills, 54% of the interventions found significant benefits, with combined interventions resulting in large effects, while FMS and/or PA only interventions yielded small effects. Similar trend was also observed in terms of numeracy, where 83% of the combined interventions found significant medium effects, while none of the FMS and/or PA only interventions did.

While these findings suggest that combined interventions are superior to FMS and/or PA interventions, it is important to note that in the vast majority of the combined interventions, the outcome measure was directly related to the contents of the intervention. Therefore, it cannot be concluded that combining cognitive tasks to FMS and/or PA is superior per se; rather it is possible that the differences stem simply from direct practice and assessing of a particular skill. In order to address this question, studies should include an additional control group that receives the same academic content without the FMS and/or PA component. In the present review, only two of the included studies were designed in accordance (Duncan et al., 2019; Callcott et al., 2015a). Although both studies demonstrated the superiority of a combined intervention to an FMS and/or PA only intervention, only one found the combined intervention to be superior to the academic control group (Duncan et al., 2019), while the other showed no difference between the two (Callcott et al., 2015a). Thus, more studies with the requisite design are needed before drawing definitive conclusions. When examining the effects of FMS and PA interventions, FMS interventions demonstrated greater effects on cognitive and academic skills compared to PA interventions (small-to-large vs. trivial-to-small effects). The superiority of FMS interventions might be explained by the fact that especially complex FMS tasks activate the same brain areas as complex cognitive tasks, and the co-activation in these brain areas is one potential explanation for the close associations between learning these skills (Diamond, 2000). In terms of PA, it is suggested that especially PA with higher duration and frequency can be effective on children's cognitive and academic skills, albeit this evidence is based on a small number of studies with weak quality (Carson et al., 2016). It is also possible that higher PA intensity plays a role in this association (Carson et al., 2016) as it was demonstrated by Chang et al. (2013) reporting superior effects of moderate-intensity in comparison to low-intensity training. However, further high-quality studies in this age group are required to confirm these findings.

An important consideration is that in four of the included studies, the significant effects were limited to within-group comparisons despite the inclusion of a control group (Bedard et al., 2020; Stein et al., 2017; Lai et al., 2020; Beck et al., 2016). This raises a general concern in regard to the validity of the presented within-group effects in the current review as well as highlights the importance of applying a control-group design for establishing intervention effects. Moreover, only four of the included studies had sufficient statistical power to detect a medium effect size (Callcott et al., 2015a; Shoval et al., 2018; Have et al., 2018; Sánchez-López et al., 2019). In part, the lack of statistical power appeared to be the consequence of the absence of a priori power calculations; which only nine of the included studies reported having conducted (Have et al., 2018; Habibi et al., 2014; Stein et al., 2017; Raney et al., 2017; Gao et al., 2019; Jarraya et al., 2019; Lai et al., 2020; Wen et al., 2018; Xiong et al., 2019a). Interestingly, however, even within the aforementioned nine studies, only one was adequately powered (Have et al., 2018) and five stated low statistical power as a limitation of the study (Gao et al., 2019; Jarraya et al., 2019; Lai et al., 2020; Wen et al., 2018; Xiong et al., 2019a). This suggests that the concept of statistical power, similarly to other modern research fields (Dumas-Mallet et al., 2017), remains widely under-appreciated – and perhaps misunderstood – in the present field of research. This is an important consideration as underpowered studies may indicate substantially inflated effects as well as lead to non-significant results when a true effect is observed (La Caze and Duffull, 2011). Nonetheless, considering the unequivocal inability of the studies to detect small intervention effects, it is probable that the observed average effects in the present review are inflated compared to the true effects. Therefore, in order to improve the validity of scientific research, authors are strongly encouraged to ensure adequate statistical power by conducting a priori power analyses.

Finally, most of the included intervention studies had a weak methodological quality and thus the results have to be dealt with caution. The finding is in line with previous systematic reviews (Carson et al., 2016). In general, low ratings resulted from inadequate reporting regarding several fundamental components such as relevant confounders, blinding, validity and reliability of the outcome measures, the agreement percentage of the selected participants as well as withdrawals and drop-outs. Previous systematic reviews with older children have found similar insufficiencies in reporting practices (Watson et al., 2017). In order to improve the quality of research evidence, adherence to published reporting guidelines for intervention studies (e.g., CONSORT 2010 Statement for RCT; Schulz et al., 2010) is highly encouraged for future studies.

4.1. Study limitations and strengths

The present systematic review had numerous strengths. FMS and PA interventions as well as combined interventions were examined, included studies were conducted in 17 countries representing almost every continent, and all studies, except one, included both girls and boys. In addition, effect size calculations allowed the quantification and comparison of intervention effects across studies. However, this review also included some limitations. First, only studies published in English were included; making it possible that some relevant studies were missed. Further, the target population was limited to typically developing children and thus, future research should also cover children with special needs.

5. Conclusions

Collectively, these results indicate that it may be possible to support typically developing preschoolers' cognitive and academic learning with FMS and PA interventions. When comparing FMS and PA only interventions to combined interventions, the latter appeared to have larger effects. However, due to the lack of essential control groups, this notion remains unelucidated. In addition, FMS interventions were found to have larger effects compared to PA interventions. Nonetheless, most of the analysed studies had low methodological quality and thus these results should be considered with caution. This highlights the need for high quality intervention studies with representative and sufficient sample sizes, control groups, as well as validated data collection tools and adequate reporting practices.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ypmed.2021.106948.

Declaration of Competing Interest

None. Funding sources had no involvement in the preparation, conducting nor in the decision to submit the article for publication.

Acknowledgments

This work was supported by the Ministry of Education and Culture [grant number 48825]; and the Finnish Cultural Foundation [Huhtamäen rahasto 2019, no grant number available].

References

Armijo-Olivo, S., Stiles, C.R., Hagen, N.A., Biondo, P.D., Cummings, G.G., 2012. Assessment of study quality for systematic reviews: a comparison of the Cochrane collaboration risk of Bias tool and the effective public health practice project quality assessment tool: methodological research. J. Eval. Clin. Pract. 18 (1), 12–18. https://doi.org/10.1111/j.1365-2753.2010.01516.x.

Bala, G., Krneta, Z., Drid, P., 2013. The effects of kinesiological activity on motor and intellectual functioning of children in relation to their physical constitution at birth. Colleg. Antropol. 37 (Suppl. 2), 161–169.

Barnard, M., Deventer, K.V., Oswald, M.M., 2014. The role of active teaching programmes in academic skills enhancement of grade 12 learners in the Stellenbosch region. S. Afr. J. Res. Sport Phys. Educ. Recreat. 36 (2), 1–14. https://doi.org/10.4314/sajrs.v36i2.

Bart, O., Hajami, D., Bar-Haim, Y., 2007. Predicting school adjustment from motor abilities in kindergarten. Infant Child Dev. 16 (6), 597–615. https://doi.org/10.1002/icd.514.

Battaglia, G., Alesi, M., Tabacchi, G., Palma, A., Bellafiore, M., 2019. The development of motor and pre-literacy skills by a physical education program in preschool children: a non-randomized pilot trial. Front. Psychol. 9 https://doi.org/10.3389/fpsyg.2018.02694.

Beck, M.M., Lind, R.R., Geertsen, S.S., Ritz, C., Lundbye-Jensen, J., Wienecke, J., 2016. Motor-enriched learning activities can improve mathematical performance in preadolescent children. Front. Hum. Neurosci. 10 https://doi.org/10.3389/fnhum.2016.00645.

Bedard, C., Bremer, E., Campbell, W., Cairney, J., 2017. A quasi-experimental study of a movement and Preliteracy program for 3- and 4-year-old children. Front. Pediatr. 5 https://doi.org/10.3389/fped.2017.00094.

Bedard, C., Bremer, E., Cairney, J., 2020. Evaluation of the move 2 learn program, a community-based movement and pre-literacy intervention for young children. Phys. Educ. Sport Pedagog. 25 (1), 101–117. https://doi.org/10.1080/17408989.2019.1690645.

Botha, S., Africa, E.K., 2020. The effect of a perceptual-motor intervention on the relationship between motor proficiency and letter knowledge. Early Childhood Educ. J. 48 (6), 727–737. https://doi.org/10.1007/s10643-020-01034-8.

Burger, K., 2010. How does early childhood care and education affect cognitive development? An international review of the effects of early interventions for children from different social backgrounds. Early Child. Res. Q. 25 (2), 140–165. https://doi.org/10.1016/j.ecresq.2009.11.001.

Byrt, T., 1996. How good is that agreement? Epidemiology 7 (5), 561. https://journals.lww.com/epidem/Citation/1996/09000/How_Good_Is_That_Agreement_.30.aspx.

Carson, V., Kuzik, N., Hunter, S., Wiebe, S.A., Spence, J.C., Friedman, A., Tremblay, M.S., Slater, L.G., Hinkley, T., 2015. Systematic review of sedentary behavior and cognitive development in early childhood. Prev. Med. 78, 115–122. https://doi.org/10.1016/j.ypmed.2015.07.016.

Callcott, D., Hammond, L., Hill, S., 2015a. The synergistic effect of teaching a combined explicit movement and phonological awareness program to preschool aged students. Early Childhood Educ. J. 43 (3), 201–211. https://doi.org/10.1007/s10643-014-0652-7.

Callcott, D, Hammond, L, Hill, S, 2015b. The synergistic effect of teaching a combined explicit movement and phonological awareness program to preschool aged students. Early Childhood Educ. J. 43 (3), 201–211. https://doi.org/10.1007/s10643-014-0652-7.

Carson, V., Hunter, S., Kuzik, N., Wiebe, S.A., Spence, J.C., Friedman, A., Tremblay, M.S., Slater, L., Hinkley, T., 2016. Systematic review of physical activity and cognitive development in early childhood. J. Sci. Med. Sport 19 (7), 573–578. https://doi.org/10.1016/j.jsams.2015.07.011.

Carson, V., Lee, E.-Y., Hewitt, L., Jennings, C., Hunter, S., Kuzik, N., Stearns, J.A., Unrau, S.P., Poitras, V.J., Gray, C., Adamo, K.B., Janssen, I., Okely, A.D., Spence, J. C., Timmons, B.W., Sampson, M., Tremblay, M.S., 2017. Systematic review of the relationships between physical activity and health indicators in the early years (0-4 years). BMC Public Health 17 (5), 854. https://doi.org/10.1186/s12889-017-4860-0.

Caspersen, C.J., Powell, K.E., Christenson, G.M., 1985. Physical activity, exercise, and physical fitness: definitions and distinctions for health-related research. Public Health Rep. 100 (2), 126–131. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1424733/.

Chang, Y.-K., Tsai, Y.-J., Chen, T.-T., Hung, T.-M., 2013. The impacts of coordinative exercise on executive function in kindergarten children: an ERP study. Exp. Brain Res. 225 (2), 187–196. https://doi.org/10.1007/s00221-012-3360-9.

Cohen, J., 1988. Statistical Power Analysis for the Behavioral Sciences, Chapter 1, the Concepts of Power Analyses, 2nd ed. Lawrence Erlbaum Associates, United States of America.

Diamond, A., 2000. Close interrelation of motor development and cognitive development and of the cerebellum and prefrontal cortex. Child Dev. 71 (1), 44–56. https://doi.org/10.1111/1467-8624.00117.

Dumas-Mallet, E., Button, K., Boraud, T., Gonon, F., Munafò, M., 2017. Low statistical power in biomedical science: a review of three human research domains. R. Soc. Open Sci. 4 (2), 160254 https://doi.org/10.1098/rsos.160254.

Duncan, G.J., Dowsett, C.J., Claessens, A., Magnuson, K., Huston, A.C., Klebanov, P., Pagani, L.S., Feinstein, L., Engel, M., Brooks-Gunn, J., Sexton, H., Duckworth, K., Japel, C., 2007. School readiness and later achievement. Dev. Psychol. 43 (6), 1428–1446. https://doi.org/10.1037/0012-1649.43.6.1428.

Duncan, M., Cunningham, A., Eyre, E., 2019. A combined movement and story-telling intervention enhances motor competence and language ability in pre-schoolers to a greater extent than movement or story-telling alone. Eur. Phys. Educ. Rev. 25 (1), 221–235. https://doi.org/10.1177/1356336X17715772.

Ellemberg, D., St-Louis-Deschênes, M., 2010. The effect of acute physical exercise on cognitive function during development. Psychol. Sport Exerc. 11 (2), 122–126. https://doi.org/10.1016/j.psychsport.2009.09.006.

Faul, F., Erdfelder, E., Lang, A.-G., Buchner, A., 2007. G*power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behav. Res. Methods 39 (2), 175–191. https://doi.org/10.3758/bf03193146.

Fischer, U., Moeller, K., Bientzle, M., Cress, U., Nuerk, H.-C., 2011. Sensori-motor spatial training of number magnitude representation. Psychon. Bull. Rev. 18 (1), 177–183. https://doi.org/10.3758/s13423-010-0031-3.

Fisher, A., Boyle, J.M., Paton, J.Y., Tomporowski, P., Watson, C., McColl, J.H., Reilly, J. J., 2011. Effects of a physical education intervention on cognitive function in young children: randomized controlled pilot study. BMC Pediatr. 11 (1), 97. https://doi.org/10.1186/1471-2431-11-97.

Gallahue, D.L., Ozmun, J.C., 2002. Understanding Motor Development: Infants, Children, Adolescents, Adults. McGraw-Hill.

Gao, Z., Lee, J.E., Zeng, N., Pope, Z.C., Zhang, Y., Li, X., 2019. Home-based exergaming on Preschoolers' energy expenditure, cardiovascular fitness, body mass index and cognitive flexibility: a randomized controlled trial. J. Clin. Med. 8 (10), 1745. https://doi.org/10.3390/jcm8101745.

Habibi, A., Ilari, B., Crimi, K., Metke, M., Kaplan, J.T., Joshi, A.A., Leahy, R.M., Shattuck, D.W., Choi, S.Y., Haldar, J.P., Ficek, B., Damasio, A., Damasio, H., 2014. An equal start: absence of group differences in cognitive, social, and neural measures prior to music or sports training in children. Front. Hum. Neurosci. 8 https://doi.org/10.3389/fnhum.2014.00690.

Have, M., Nielsen, J.H., Ernst, M.T., Gejl, A.K., Fredens, K., Grøntved, A., Kristensen, P.L., 2018. Classroom-based physical activity improves children's math achievement – a randomized controlled trial. PLoS One 13 (12), e0208787. https://doi.org/10.1371/journal.pone.0208787.

Hinkley, T., Salmon, J., Okely, A.D., Crawford, D., Hesketh, K., 2012. Preschoolers' physical activity, screen time, and compliance with recommendations. Med. Sci. Sports Exerc. 44 (3), 458–465. https://doi.org/10.1249/MSS.0b013e318233763b.

Jarraya, S., Wagner, M., Jarraya, M., Engel, F.A., 2019. 12 weeks of kindergarten-based yoga practice increases visual attention, visual-motor precision and decreases behavior of inattention and hyperactivity in 5-year-old children. Front. Psychol. 10 https://doi.org/10.3389/fpsyg.2019.00796.

Jones, D., Innerd, A., Giles, E.L., Azevedo, L.B.D., 2020. Association between fundamental motor skills and physical activity in the early years: a systematic review and meta-analysis. J. Sport Health Sci. 9 (6), 542–552. https://doi.org/10.1016/j.jshs.2020.03.001.

La Caze, A., Duffull, S., 2011. Estimating risk from underpowered, but statistically significant, studies: was APPROVe on TARGET? J. Clin. Pharm. Ther. 36 (6), 637–641. https://doi.org/10.1111/j.1365-2710.2010.01222.x.

Lai, Y., Wang, Z., Yue, G.H., Jiang, C., 2020. Determining whether tennis benefits the updating function in young children: a functional near-infrared spectroscopy study. Appl. Sci. 10 (1), 407. https://doi.org/10.3390/app10010407.

Lee, J., Zhang, T., Chu, T.L., (Alan), Gu, X., & Zhu, P., 2020. Effects of a fundamental motor skill-based afterschool program on Children's physical and cognitive health outcomes. Int. J. Environ. Res. Public Health 17 (3). https://doi.org/10.3390/ijerph17030733.

Logan, S.W., Ross, S.M., Chee, K., Stodden, D.F., Robinson, L.E., 2018. Fundamental motor skills: a systematic review of terminology. J. Sports Sci. 36 (7), 781–796. https://doi.org/10.1080/02640414.2017.1340660.

Malina, R.M., Bouchard, C., Bar-Or, O., 2004. Growth, Maturation, and Physical Activity. Human Kinetics.

Mavilidi, M.-F., Okely, A.D., Chandler, P., Cliff, D.P., Paas, F., 2015. Effects of integrated physical exercises and gestures on preschool Children's foreign language vocabulary learning. Educ. Psychol. Rev. 27 (3), 413–426. https://doi.org/10.1007/s10648-015-9337-z.

Mavilidi, M.-F., Okely, A., Chandler, P., Louise Domazet, S., Paas, F., 2018a. Immediate and delayed effects of integrating physical activity into preschool children's learning of numeracy skills. J. Exp. Child Psychol. 166, 502–519. https://doi.org/10.1016/j.jecp.2017.09.009.

Mavilidi, M.-F., Okely, A., Chandler, P., Louise Domazet, S., Paas, F., 2018b. Immediate and delayed effects of integrating physical activity into preschool children's learning of numeracy skills. J. Exp. Child Psychol. 166, 502–519. https://doi.org/10.1016/j.jecp.2017.09.009.

Mavilidi, M.-F., Okely, A.D., Chandler, P., Paas, F., 2016a. Infusing physical activities into the classroom: effects on preschool Children's geography learning. Mind Brain Educ. 10 (4), 256–263. https://doi.org/10.1111/mbe.12131.

Mavilidi, M.-F., Okely, A.D., Chandler, P., Paas, F., 2016b. Infusing physical activities into the classroom: Effects on preschool children's geography learning. Mind Brain Educ. 10 (4), 256–263. https://doi.org/10.1111/mbe.12131.

Mavilidi, M.-F., Okely, A.D., Chandler, P., Paas, F., 2017a. Effects of integrating physical activities into a science lesson on preschool Children's learning and enjoyment. Appl. Cogn. Psychol. 31 (3), 281–290. https://doi.org/10.1002/acp.3325.

Mavilidi, M.-F., Okely, A.D., Chandler, P., Paas, F., 2017b. Effects of integrating physical activities into a science lesson on preschool children's learning and enjoyment. Appl. Cong. Psychol. 31 (3), 281–290. https://doi.org/10.1002/acp.3325.

Melnyk, B.M., Fineout-Overholt, E., 2005. Evidence-Based Practice in Nursing & Healthcare: A Guide to Best Practice. Lippincott Williams & Wilkins, Philadelphia, p. 29.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D.G., 2009. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. BMJ 339, b2535. https://doi.org/10.1136/bmj.b2535.

National Association for Sport and Physical Education, 2002. Active Start: A Statement of Physical Activity Guidelines for Children Birth to 5 Years. AAHPERD Publications, Oxon Hill, MD.

National Collaborating Centre for Methods and Tools, 2008. Quality Assessment Tool for Quantitative Studies. McMaster University, Hamilton, ON.

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. Prev. Med. 72, 116–125. https://doi.org/10.1016/j.ypmed.2014.12.027.

Padial-Ruz, R., García-Molina, R., Puga-González, E., 2019. Effectiveness of a motor intervention program on motivation and learning of English vocabulary in preschoolers: a pilot study. Behav. Sci. 9 (8), 84. https://doi.org/10.3390/bs9080084.

Pienaar, Anita E., Esme, Van Rensburg, Adele, Smit, 2011. Effect of a Kinderkinetics programme on components of children's perceptual-motor and cognitive functioning. S. Afr. J. Res. Sport Phys. Educ. Recreat. 33 (3), 113–128. https://doi.org/10.10520/EJC108959.

Planinšec, J., 2002. Relations between motor and cognitive dimensions of preschool girls and boys. Percept. Mot. Skills 94, 415–423. https://doi.org/10.2466/pms.2002.94.2.415.

Ramah, N., 2014. The effects of square-stepping exercises on cognitive skills for kindergarten age children. Ovidius Univ. Ann. Ser. Phys. Educ. Sport/Sci. Mov. Health 14 (1), 126–131. https://go.gale.com/ps/i.do?p=AONE&sw=w&issn=2285777X&v=2.1&it=r&id=GALE%7CA417737975&sid=googleScholar&linkaccess=abs.

Raney, M., Henriksen, A., Minton, J., 2017. Impact of short duration health & science energizers in the elementary school classroom. Cogent Educ. 4 (1), 1399969. https://doi.org/10.1080/2331186X.2017.1399969.

Rasmussen, M., Laumann, K., 2013. The academic and psychological benefits of exercise in healthy children and adolescents. Eur. J. Psychol. Educ. 28 (3), 945–962. https://doi.org/10.1007/s10212-012-0148-z.

Roth, K., Ruf, K., Obinger, M., Mauer, S., Ahnert, J., Schneider, W., Graf, C., Hebestreit, H., 2010. Is there a secular decline in motor skills in preschool children? Scand. J. Med. Sci. Sports 20 (4), 670–678. https://doi.org/10.1111/j.1600-0838.2009.00982.x.

Ruiz-Hermosa, A., Mota, J., Díez-Fernández, A., Martínez-Vizcaíno, V., Redondo-Tébar, A., Sánchez-López, M., 2020. Relationship between weight status and cognition in children: a mediation analysis of physical fitness components. J. Sports Sci. 38 (1), 13–20. https://doi.org/10.1080/02640414.2019.1676538.

Sánchez-López, M., Cavero-Redondo, I., Álvarez-Bueno, C., Ruiz-Hermosa, A., Pozuelo-Carrascosa, D.P., Díez-Fernández, A., Gutierrez-Díaz Del Campo, D., Pardo-Guijarro, M.J., Martínez-Vizcaíno, V., 2019. Impact of a multicomponent physical activity intervention on cognitive performance: the MOVI-KIDS study. Scand. J. Med. Sci. Sports 29 (5), 766–775. https://doi.org/10.1111/sms.13383.

Schulz, K.F., Altman, D.G., Moher, D., 2010. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. BMJ 340, c332. https://doi.org/10.1136/bmj.c332.

Schulz, K.F., Grimes, D.A., 2005. Sample size calculations in randomised trials: mandatory and mystical. The Lancet 365 (9467), 1348–1353. https://doi.org/10.1016/S0140-6736(05)61034-3.

Shoval, E., Sharir, T., Arnon, M., Tenenbaum, G., 2018. The effect of integrating movement into the learning environment of kindergarten children on their academic achievements. Early Childhood Educ. J. 46 (3), 355–364. https://doi.org/10.1007/s10643-017-0870-x.

St. Laurent, C.W.C.W., Burkart, S., Alhassan, S., 2018. Effect of a School-Based Physical Activity Intervention on Number and Letter Recognition in Preschoolers. Int J. Exerc. Sci. 11 (5), 168–178.

Stein, M., Auerswald, M., Ebersbach, M., 2017. Relationships between motor and executive functions and the effect of an acute coordinative intervention on executive functions in kindergartners. Front. Psychol. 8 https://doi.org/10.3389/fpsyg.2017.00859.

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. https://doi.org/10.1080/00336297.2008.10483582.

Thomas, B.H., Ciliska, D., Dobbins, M., Micucci, S., 2004. A process for systematically reviewing the literature: providing the research evidence for public health nursing interventions. Worldviews Evid.-Based Nurs. 1 (3), 176–184. https://doi.org/10.1111/j.1524-475X.2004.04006.x.

Tomporowski, P.D., Lambourne, K., Okumura, M.S., 2011. Physical activity interventions and children's mental function: an introduction and overview. Prev. Med. 52, S3–S9. https://doi.org/10.1016/j.ypmed.2011.01.028.

Tonge, K.L., Jones, R.A., Okely, A.D., 2016. Correlates of children's objectively measured physical activity and sedentary behavior in early childhood education and care services: a systematic review. Prev. Med. 89, 129–139. https://doi.org/10.1016/j.ypmed.2016.05.019.

Toumpaniari, K., Loyens, S., Mavilidi, M.-F., Paas, F., 2015. Preschool Children's foreign language vocabulary learning by embodying words through physical activity and gesturing. Educ. Psychol. Rev. 27 (3), 445–456. https://doi.org/10.1007/s10648-015-9316-4.

Tucker, P., 2008. The physical activity levels of preschool-aged children: a systematic review. Early Child. Res. Q. 23 (4), 547–558. https://doi.org/10.1016/j.ecresq.2008.08.005.

Watson, A., Timperio, A., Brown, H., Best, K., Hesketh, K.D., 2017. Effect of classroom-based physical activity interventions on academic and physical activity outcomes: a systematic review and meta-analysis. Int. J. Behav. Nutr. Phys. Activity 14 (1), 114. https://doi.org/10.1186/s12966-017-0569-9.

Wen, X., Zhang, Y., Gao, Z., Zhao, W., Jie, J., Bao, L., 2018. Effect of mini-trampoline physical activity on executive functions in preschool children. Biomed. Res. Int. 2018, 2712803. https://doi.org/10.1155/2018/2712803.

Xiong, S., Li, X., Tao, K., 2017. Effects of structured physical activity program on Chinese young Children's executive functions and perceived physical competence in a day care center [research article]. Biomed. Res. Int. https://doi.org/10.1155/2017/5635070.

Xiong, S., Zhang, P., Gao, Z., 2019a. Effects of exergaming on Preschoolers' executive functions and perceived competence: a pilot randomized trial. J. Clin. Med. 8 (4), 469. https://doi.org/10.3390/jcm8040469.

Xiong, S., Zhang, P., Gao, Z., 2019b. Effects of exergaming on preschoolers' executive functions and perceived competence: a pilot randomized trial. J. Clin. Med. 8 (4), 469. https://doi.org/10.3390/jcm8040469.

Zeng, N., Ayyub, M., Sun, H., Wen, X., Xiang, P., Gao, Z., 2017. Effects of physical activity on motor skills and cognitive development in early childhood: a systematic review [review article]. Biomed. Res. Int. https://doi.org/10.1155/2017/2760716.