Table 1. Descriptives and methodological quality of the included intervention studies (sorted from strong to weak, in alphabetical order)

Source	Study design	Sample characteristics (baseline)	Confounders	Blinding	AP measure	CP measure	Analytical sample	Quality	Main results
Ahamed et al., 2007	Cluster RCT, Action Schools! BC (AS! BC) Intervention + Control, Canada Duration: 16 months Type: classroom-based PA Bout duration: 15 min Frequency: 5 days/week Intensity: can't tell Control: 2 x 40 min of regular PE	Age: 9-11 years N = 396	Accounted for gender, ethnicity, PA score and baseline differences in AP	Assessors: yes Participants: can't tell	Canadian Achievement Test (e.g. mathematics, reading and language)		>70% ~50% female	Strong	PA: intervention > control (approximately 47 min/week) AP: intervention = control
Chaya et al., 2012	RCT, Intervention, India Duration: 3 months Type: 1) Yoga	Age: 7-9 years N = 200 ~49% female	Accounted for gender, age, BMI, SES and baseline CP	Assessors: yes Participants: can't tell		Malin's intelligence scale for Indian children	Cohort retention: >90% Analytical sample: >90%	Strong	PA: no information CP: yoga = PA group [all children showed improved

	Bout duration: 45 min Frequency: 6 days/week Intensity: can't tell 2) PA group (passive stretching and aerobic exercise) Bout duration: 45 min Frequency: 6 days/week Intensity: can't tell								performance on comprehension, arithmetic, vocabulary and object assembly scores at 3 (posttest) and 6 months (baseline - follow-up) compared to the baseline scores]
Donnelly et al., 2009	Cluster RCT, Physical Activity Across the Curriculum (PAPC) Intervention + Control, US Duration: 3 years Type: physically active academic lessons Bout duration: 10 min Frequency: 5 days/week,	Age: 7-9 N = 1527	Accounted for gender, age, BMI, AP and daily PA	Assessors: yes Participants: can't tell	Wechsler Individual Achievement Test 2nd Edition assessing mathematics reading, writing and oral language skills		>90 %** ~57% female **refers to % of participants that completed the intervention; outcomes were	Strong	PA: intervention > control (13% more active) AP: intervention > control for composite academic scores of reading, mathematics and spelling

	amounting to 90 min/week Intensity: MVPA Control: Regular academic lessons and regular PE (60 min/week)						measured in a subsample of 454 participants		
Ericsson, 2008	CT, Bunkeflo project, Sweden Duration: 3 years Type: Daily PE + 60 min adapted motor training/week (if needed) Bout duration: 45 min Frequency: 5 days/week Intensity: can't tell Control: 2 PE lessons (45 min each) per week	Age: 7-9 years N = 251	Accounted for age, parental education level, baseline CP and AP and leisure PA	Assessors: can't tell Participants: can't tell	Standardized grades for literacy (Swedish) and mathematics; qualification for upper secondary school (teacher report)	Conner's questionnaire measuring teacher's and parent's conceptions of children's attention ability and impulse control	100%	Strong	PA: no information intervention > control for motor skills improved (year 1 and year 3) CP: intervention > control for attention measures (year 1 and year 2, not in year 3) AP: Literacy and Intervention > control for mathematics (year 2)

Gao et al., 2013	CT, Dance Dance Revolution (DDR) Intervention, US Duration: 2 years Type: aerobic activities (15 min) and dance activities (15 min) Bout duration: 30 min Frequency: 3 days/week Intensity: MVPA Control: Unstructured recess	Age: 8-11 years N = 268	Accounted for gender, age, BMI, fitness and baseline AP	Assessors: can't tell Participants: can't tell	Standardized school grades on reading and mathematics		Cohort retention: >60% Analytical sample: >75% ~42% female** **analytical sample	Strong	PA: Fitness: intervention > control AP: intervention> control for mathematics scores (year 1 and year 2) Intervention = control for reading
Hillman et al., 2014	RCT, FITKids Intervention, US Duration: 9 months Type: aerobic + 45-55 min skill theme (not exercise) Bout duration: 70 min Frequency: 5 days/week Intensity: light to MVPA	Age: 7-9 N = 221 ~ 46% female	Accounted for age, pubertal timing fitness, BMI and IQ	Assessors: yes Participants: can't tell		Modified flanker task and color-shape switch task	Cohort retention: >90% Analytical sample: 100%	Strong	PA: intervention > control for fitness; mean heart rate during the PA section of the intervention was 137 beats/minute CP: Attentional inhibition (flanker task): - intervention > control for accuracy

	Control: Regular PE curriculum (can't tell)								scores and P3 amplitude and P3 latency; - intervention = control for response time: Cognitive flexibility (switch-task): - intervention > control for performance accuracy and P3 amplitude (heterogeneous trials); - intervention = control for performance accuracy and P3 amplitude (homogenous trials)
Kirk et al., 2014	CT, Intervention, US Duration: 6 months Type: academically active lessons Bout duration: 2 x 15 min Frequency: 5 days/week Intensity: light to MVPA	Age: 3-4 years N =72 53% female	Accounted for gender, age, weight and baseline AP	Assessors: can't tell Participants: can't tell	Growth and Development Indicators (picture naming, rhyming, alliteration)		100%	Strong	PA: intervention > control for PA level AP: - intervention > control for scores on picture naming task and alliteration administration (3 and 6 months);

	Control: Similar academic lessons without the PA component								- intervention = control for rhyming administration
Resaland et al., 2016	RCT, Norway Duration: 7 months Intervention with 3 components (a/b/c) Frequency/Bout duration: a) 90 min/week physically active educational lessons in school playground b) 5 min/day PA breaks during classroom lessons c) 10 min/day PA homework Intensity: 25% of PA was vigorous PA Control: Regular PE/PA curriculum (135 min/week)	Age: mean 10.2 years 57 elementary schools N = 1129	Accounted for age, sex, parents' education level, weight, height, BMI, baseline AP and baseline PA	Assessors: can't tell Participants: can't tell	Academic performance numeracy, reading, English using standardized Norwegian Tests		~95%	Strong	PA: accelerometry AP: intervention = control for all AP measures; subgroup intervention > control for numeracy

Schmidt et al., 2015	RCT, intervention program; Switzerland Duration: 6 weeks Type: 1) cognitively engaging team games intervention (high demands on cognitive engagement and physical exertion)and 2) one aerobic intervention (low cognitive engagement but high physical exertion) Bout duration: 45 min Frequency: 2 days/week Intensity: MVPA Control: Standard PE consisting of low cognitive engagement and	Age: 10-12 years N = 118 55% female	Accounted for gender, age, SES, BMI, baseline PA and AP	Assessors: partially Participants: can't tell		Non-spatial n-back (updating) Flanker test (inhibition) Mixed cued flanker (shifting)	>90%	Strong	PA: Team group and aerobic intervention > control for physical exertion and physical fitness. No group differences for time spent being physically active. CP: - intervention = control for inhibition and updating - team games (1) > aerobic intervention (2) or control group, for shifting

	low physical exertion demands								
Subramania n et al., 2016	RCT, Structured PA Intervention program, India Duration: 6 months Type: Structured PA (aerobic) Bout duration: 120 min Frequency: 6 days/week Intensity: can't tell Control: Unstructured PA for the same bout, frequency and duration as the intervention group	Age: 12-17 years N = 451 44% female Both the intervention and the control group are divided in athletes and non-athletes.	Accounted for gender, age, height, weight, baseline PA and baseline CP	Assessors: yes Participants: can't tell		Two target letter cancellation test (LCT) Trail Making test A and B (TTA and TTB) Ruff Figural fluency test (RFFT)	>75%	Strong	PA: No information CP: intervention > control for LCT time, TTB and RFFT intervention = control for LCT omission, LCT commission, TTA
Telford et al., 2012	Cluster RCT, specialist-Taught Intervention, Australia Duration: 2 years Type: Specialist-Taught PE	Age: 8-9 years N = 750	Accounted for gender, age, BMI, baseline AP and baseline PA	Assessors: can't tell Participants: can't tell	Literacy and numeracy test		~83% 49% female	Strong	PA: intervention = control AP: intervention > control for numeracy

	Bout duration: 45-50 min Frequency: 2 days/week Intensity: MVPA + 50-60 min of PE per week delivered by the classroom teacher Control: 150 min per week of PE lessons taught by classroom teachers as well as sports teachers						>80%		
Beck et al., 2016	CT, 2 types of motor-enriched learning activities, Denmark Duration: 6 weeks (T0/T1) Follow-up after 8 weeks after intervention stopped (T2) Frequency: 3 days/week	Age: mean 7.5 years N = 165 Intervention (1) GMM 55% male N = 55 Intervention (2) FMM: 57% male, N = 53 Control 49% male	Accounted for age, gender, bilingualism, BMI, fitness, gross motor skills, fine motor skills, executive functions, visuo-spatial short-term memory and	Assessors: can't tell Participants: can't tell	Mathematical achievement test (arithmetics and geometry)		>80%	Moderate	PA: Children attending classes: all groups >90% objectively measured HR zone in subgroup: MVPA HR GMM > FMM/CO AP: *Total group* GMM > FMM (T0 – T1) *Normal learners*

| | Bout duration: 60 min
Intensity: can't tell

Intervention (1)
Gross Motor Math
GMM: inter-limb gross motor movements, involving a large range of movements while solving mathematical problems

Intervention (2)
Fine Motor Math
FMM: using fine motor activity to manipulate LEGO bricks supporting mathematical principles and procedures.

Control:
Conventional math teaching | N = 57 | phonological short-term memory | | | | | GMM>FMM/CON (T0-T1)
GMM>CON (T0—T2) |

Bunketorp-Kall et al., 2015	Quasi-experimental design, ''School-in Motion'' intervention program, Sweden Duration: 3 years Type: aerobic Bout duration: 30-45 min Frequency: 2 days/week Intensity: can't tell	Age: 5-12 years N = 545	Accounted for gender, age, SES, baseline AP and PA	Assessors: can't tell Participants: can't tell	National test results in mathematics, Swedish and English ("pass" or "fail")		>60% ~48% female	Moderate	PA: intervention > control AP: Intervention > control for passing national tests in Swedish and mathematics in girls (not boys); Intervention = control for English grades.
Butzer et al., 2015	CT, "Kripalu Yoga in the Schools" Intervention, US Duration: 12 weeks Type: Yoga Bout duration: 35-40 min Frequency: 2-3 days/week Intensity: can't tell Control: 35 - 40 min of PE for 2 - 3 times per week, for a total duration of 12	Age: 14-16 years N = 115	Accounted for gender and age	Assessors: can't tell Participants: can't tell	Standardized GPA		>80% ~ 56% female	Moderate	PA: no information AP/CP: Both groups showed a downwards decline in GPA between the 1st and 4th quarter. However, the decline between the 1st (baseline) quarter and the 3rd quarter (intervention was delivered) was more pronounced for the control than the yoga group. The effect of

	weeks, as well as theory over sports								yoga practice did not last after the practice was discontinued.
Castelli et al., 2011	Pre-posttest design, part of the FITKids Intervention, US Duration: 9 months Type: PA sessions Bout duration: 70 min Frequency: 5 days/week Intensity: light to MVPA +45-55 min skill theme Control: none	Age: 8-9 years N = 59 *	Not Applicable	Assessors: yes Participants: can't tell	Wide Range Achievement Test (WRAT)	Kaufman Brief Intelligence Test (KBIT) Trail making A and B The Stroop Color-Word test	100% 44% female	Moderate	PA: not applicable AP/CP: Children who spent more time exercising above the pre-specified target heart zone performed better at tasks related to executive functions (Trial making B and Stroop Color-Word) after the intervention. Mean heart rate during PA sessions was related to performance on the Trial making A task.
Chaddock et al., 2013	RCT, FITKids Intervention, US Duration: 9 months Type: aerobic Bout duration: 70 min	Age: 8-9 N = 32 ~57% female *	Accounted for gender, age, SES, pubertal timing, physical fitness and IQ scores	Assessors: yes Participants: can't tell		Modified flanker task with 3 conditions: congruent, incongruent and No Go	~72%	Moderate	PA: intervention > control for VO$_2$max (physical fitness) CP: intervention = control

	Frequency: 5 days/week Intensity: light to MVPA +45-55 min skill theme Control: Regular PE curriculum								[Posthoc tests guided by a-priori hypothesis showed higher improvements in performance between the pretest and posttest in the intervention group, for the flanker task, measuring attention and interference control. Children in the intervention group showed decreases in right anterior prefrontal cortex activation. Children in the intervention group showed similar anterior frontal brain patterns and incongruent accuracy rates to a group of college-aged young adults]
Chang et al., 2013	CT, Coordinative exercise Intervention, Taiwan	Age: 6-7.5 years N = 26 50% female	Accounted for gender, age, height, weight, SES, baseline physical fitness	Assessors: can't tell		Eriksen flanker task	Can't tell	Moderate	PA: moderate intensity group > low intensity group for fitness

	Duration: 8 weeks Type: moderate intensity coordinative exercise Bout duration: 35 min Frequency: 2 days/week Intensity: moderate Group 2: Duration: 8 weeks Type: low intensity coordinative exercise Bout duration: 35 min Frequency: 2 days/week Intensity: low		and soccer experience	Participants: can't tell					CP: moderate intensity group = low intensity group; all children, regardless of group membership had higher accuracy and faster reaction times, particularly for the task condition requiring inhibitory control. The behavioural results were accompanied by larger P3 amplitudes and shorter P3 latencies, a neural marker of better use of attentional resources.
Costigan et al., 2016	RCT, New Zealand Duration: 8 weeks Frequency: 3 times/weeks Bout duration: 8-10 min Intensity: high intensity	Age: 14-16 years N = 65	Accounted for age, weight, height and BMI	Assessors: yes Participants: can't tell		Trail making test (visual attention, speed, scanning, speed of processing and mental	95% (Executive Function available)	Moderate	PA: Measured using heart rate monitors (average bpm). AEP: 148 RAP: 155 No information on control.

	Interventions (delivered during PE lessons (2) and lunch time (1): (1) AEP (Aerobic Exercise Program): gross motor cardiorespiratory exercises (2) RAP (Resistance and aerobic exercise program): a combination of cardiorespiratory and body weight resistance training exercises Control: usual PE and lunchtime activities					flexibility)			CP: intervention = control
Donnelly et al., 2017	RCT, US A+PAAC type: integrated PA with academic instruction. Duration: 3 years	Age: mean 8.1 years 17 schools (9 intervention	Accounted for BMI, household income, cardiovascular fitness and	Assessors: yes Participants: can't tell	Wechsler Individual Achievement Test (WIAT-III)		77%	Moderate	PA: Direct Observations (SOFIT) AP: intervention = control

	Intervention	Population		Outcome measures				Quality	Results
	Frequency: 2 times/day, 5 days/week Bout duration: 10 min Intensity: 4-5 METs Control: Regular curriculum	, 8 control schools); N = 584	baseline AP score						
Erwin et al., 2017	CT, classroom-based PA Intervention + Control, US Duration: 20 weeks Type: classroom-based PA Bout duration: 20+ min Frequency: 5 days/week Intensity: can't tell Control: 2 x 30 min PE per week and 1 x 30 min recess session per day	Age: 8-9 years N = 29	Can't tell	Assessors: can't tell Participants: can't tell	Curriculum based measurements for reading fluency and mathematics Aptitude, grades for mathematics and reading Standardized test scores and classroom behaviour		100%	Moderate	PA: no information AP: intervention > control for reading fluency and mathematics scores and for standardized reading and mathematics scores.

Fedewa et al., 2015	CT, Physically active academic lessons, Intervention, US Duration: 8 months Type: physically active academic lessons Bout duration: 20 min Frequency: 5 days/week Intensity: can't tell Control: Regular school curriculum	Age: 8-11 years N = 460 % female (can't tell)	Baseline scores AP and IQ	Assessors: can't tell Participants: can't tell	Standardized grades in reading and mathematics	Raven's Standard Progressive Matrix (SPM)	> 95%	Moderate	PA: intervention> control for steps during the period between winter and fall only (seasonal-related effects) AP: intervention > control for reading and mathematics CP: intervention = control for fluid intelligence
Fisher et al., 2011	RCT, Scotland Duration: 10 weeks Type: high intensity PE (mostly aerobic) Bout duration: 60 min Frequency: 5 days/week	Age: 6-7 years N = 71 55% female	Accounted for gender, age SES and baseline CP	Assessors: yes Participants: yes		The Cognitive Assessment System (CAS; (planning, attention, perceptual processing, memory)	Cohort retention: >90% Analytical sample: >80%	Moderate/Weak	PA: intervention > control for total PA and time spent exercising at MVPA intensity levels CP: intervention > control for error rate in a

	Intensity: 44% sedentary and 20% MVPA Control: Duration: 10 weeks Type: regular PE (balance/coordinative) Bout duration: 60 min Frequency: 2 days/week Intensity: 61% sedentary and 9% MVPA					Cambridge Neuropsychological Test Battery (Spatial Memory Span and Spatial Working Memory) Attention Network Test			spatial working memory task and scores on the cognitive problems and inattention subscale
Galotta et al., 2015	RCT, Italy Duration: 5 months Frequency: 2 days/week Bout duration: 60 min Intensity: various intensities Interventions: (1) Traditional PA (endurance, strength, flexibility)	Age: 8-11 years 3 primary schools N = 230 (total sample) Traditional N = 78	Accounted for BMI, body fat, and baseline values CP	Assessors: can't tell Participants: can't tell		D2 test of attention	67%	Moderate	PA: no information CP: intervention = control

	(2) Coordinative PA (coordination, dexterity) Control: No PA	Coordinative PA N = 83 Control N = 69							
Hagins and Rundle, 2016	RCT, US Duration: 1 academic year (9 months) Frequency: 2 days/week Bout duration: 45 min Intensity: light PA Intervention: Mindfulness and yoga-based exercises Control: PE class (weightlifting, stationary biking, fitness exercises, games (e.g. soccer/volleyball)	9th-11th grade students (public high school) N = 112	Accounted for age, gender, weight, BMI, cardiovascular fitness, race/ethnicity, socio-economic status, AP previous year and class participation	Assessors: no Participants: can't tell	GPA		100%	Moderate	- intervention = control for GPA - intervention > control for GPA after controlling for participation rate

	Measurements 1, 5 and 9 months after start of the intervention									
Kirk and Kirk, 2016	Cohort analytic, Head Start Intervention, US

Academic lessons using moderate PA

Duration: 8 months Frequency: 5 days/week Bout duration: 2 times 30 min Intensity: various intensities

Control: Measurements at baseline, 4 and 8 months | Age: mean 4.1 years

2 Preschools, 4 classes N = 56

Intervention N = 40

Control N = 16 | Accounted for baseline values AP, gender, age and BMI | Assessors: no

Participants: can't tell | Literacy skills (picture naming, alliteration, rhyming) using early literacy individual growth and development indicators | | 96%

Intervention: N =39

Control: N =15 | Moderate | PA: Direct Observations (SOFIT) of fitness instruction time, teacher report

AP: - intervention = control for picture naming - intervention > control for rhyming and alliteration after 8 months |
| **Koutsandré ou et al., 2016** | CT, intervention program, Germany

Duration: 10 weeks after school program | Age: 9-10 years N = 99

55% female | Accounted for gender, age, BMI, pubertal status, physical fitness and baseline CP | Assessors: can't tell

Participants: can't tell | | The auditory Letter Digit Span | >70% | Moderate | PA: - heart rate: motor coordination intervention and aerobic intervention > control, - mean heart rate: aerobic > motor |

	Type: one motor demanding exercise (coordination) intervention and one aerobic intervention Bout duration: 45 min Frequency: 3 days/week Intensity: MVPA Control: regular PE lessons, 3 x 45 min per week + assisted homework sessions								CP: - pre-post differences in working-memory performance: motor coordination and aerobic intervention > control - posttest performance: motor coordination intervention > control
Kvalo et al., 2017	RCT, Norway Active School Duration: 10 months Intervention *Regular PE/PA* 3 x 45 min *Physically active academic lessons* 2 x 45 min *Physically active homework* 5 x 10 min	Age: 9-10 years 9 primary schools N =679	Accounted for number of children participating in extracurricular sports, living conditions surrounding school area, size, participation in PA leader program, gender, BMI,	Assessors: no Participants: can't tell		Executive functions (1) attention/inhibition/self-control/mental speed using the Stroop Golden color-word test (word reading/color naming/interference);	>60%	Moderate	PA: Min (self-reported by schools) intervention vs control Regular PE: 80 vs 60 PA: 68 vs 44 Physically active academic lessons: 78 Physically active homework: 39 Physical activity during recess: 26 CP:

	Physical activity during recess 5 x 10 min In total 325 min/week Intensity: various intensities Control Regular PE/PA (135 min/week)		waist circumference and baseline scores CP			(2) verbal semantic fluency; (3) working memory using the intelligence test WAIS-IV (digit span forward/backward); (4) attention, psychomotor execution, mental flexibility using the Trail Making Test Composite score of (1)-(4)			intervention = control for all the tests
Mavilidi et al., 2016	RCT, Australia Duration: 1 week Frequency: 3 times/8 days Bout duration: 10 min	Age: mean 4.9 years 8 childcare centers	Accounted for pre-test scores	Assessors: no Participants: can't tell	Knowledge test, directly related to the content of the interventions		97% Intervention (1) N = 28; Intervention (2) N =	Moderate	PA: (counts/minute and MVPA) - unintegrated > integrated - Integrated/ unintegrated > control

								29 Control N = 30		AP: intervention (integrated) > control intervention (unintegrated condition) = control
Mavilidi et al., 2017	RCT, Australia Duration: 4 weeks Frequency: 1/week Bout duration: 10 min Intensity: can't tell Interventions: (1) Intervention – integrated: performing physical activities related to the learning task	Age: mean 4.9 years 7 childcare centers N = 90 (female N = 45, male N = 45)	Accounted for age, gender, ethnicity, socio-economic status and baseline score CP	Assessors: no Participants: can't tell	Knowledge test, directly related to the content of the interventions 1) Free-recall test 2) Cued-recall test		91%	Moderate		PA: Measured using accelerometers: MVPA and average activity counts/minute AP: Children in de integrated condition Children performed better in the immediate post-test compared to the delayed post-test (6 weeks later)

	(2) Intervention - non-integrated: performing physical activities not related to the learning task Control: no physical activities involved in the learning task Measurements at baseline, 4 weeks, 10 weeks								
McClelland et al., 2015	CT, Move4words Intervention, UK Duration: 3 months Type: aerobic and perceptual-motoric Bout duration: 20 min Frequency: 5 days/week Intensity: can't tell Control: Regular PE	Age: 10-11 years N = 1955	Accounted for age and baseline AP	Assessors: can't tell Participants: can't tell	National examinations (KS2 SATs, math and language) – calculated as % of participants who exceeded level 4		>80%	Moderate/Weak	PA: no information AP: intervention > control for national math and English scores (79% of the children in the intervention group passed the exam, with a 20% increase compared to 3 years before the intervention; 66% of the children in the control group reached

								this target, that is a 7% increase from baseline measures)
	One-group time-series trial, Move4words Intervention, UK Duration: 3 months Type: aerobic and perceptual-motoric Bout duration: 20 min Frequency: 5 days/week Intensity: can't tell Control: Regular PE	Age: 7-10 years N = 113	Accounted for age and baseline AP		National curriculum levels in reading, writing and math		>80%	Moderate/Weak	PA: no information AP: 63% increase in reading rate, 88% increase in math progress and 19% increase in writing progress rate in the intervention group.

	Cohort-analytic, Move4words Intervention, UK Duration: 3 months Type: aerobic and perceptual-motoric Bout duration: 20 min Frequency: 5 days/week Intensity: can't tell Control: Normal PE lessons, can't tell the duration or intensity	Age: 7-10 years N = 51	Accounted for age and baseline AP		National curriculum levels in reading		Can't tell	Moderate/Weak	PA: no information AP: intervention > control for reading scores; participants in the intervention group were one year ahead in terms of reading scores than the control group.
van der Niet et al., 2016	Quasi-experimental design, additional PA during lunch time + Control, Netherlands Duration: 5.5 months Type: aerobic exercise and cognitively engaging activities	Age: 8-12 N = 112 50% female	Accounted for gender, age, SES, BMI, baseline PA and CP	Assessors: yes Participants: no		Golden version of the Stroop test, Visual Memory Span test, the Digit Span test, Trail making test and the Tower of London	>85%	Moderate	PA: intervention = control for physical fitness measures CP: - intervention > control for stroop test and digit span test scores;

	Bout duration: 30 min Frequency: 2 days/week Intensity: MVPA Control: standard 2 x PE per week								- intervention = control for remaining CP measures
Pesce et al., 2013	CT, Intervention, Italy Duration: 6 months Type: 1) specialist-led PE (G2) 2) specialist-led PE (G3), including cognitively engaging PE lessons Bout duration: 60 min Frequency: 1 day/week Intensity: MVPA Control: generalist-led PE (G1)	Age: 3-5 and 6-10 years N = 530	Accounted for gender, age, SES, baseline PA and baseline CP	Assessors: can't tell Participants: can't tell		Cognitive Assessment System, assessing planning and attention	47% (5-10 years) ~49% female	Moderate	PA: G3 > G1 and G2, for physically active time and time spent in MVPA intensity (G1 (12 min), G2 (23 min) and G3 (15 min)) CP: G3 > G1 and G2, typically developing children in the specialist lead cognitively engaging PE group showed more pronounced improvement in the total attention scale, in particular in the receptive attention subscale

| Pesce et al., 2016a | CT, Intervention, Italy

Duration: 6 months
Type: aerobic, coordinative and cognitively engaging
Bout duration: 60 min
Frequency: 1 day/week
Intensity: MVPA

Control:
1 hour of PE per week, delivered by the classroom teacher | Age: 5-10 years
N = 920

~50% female | Accounted for gender, age, BMI, weight, spontaneous outdoor play and baseline measures (motor and CP) | Assessors: yes

Participants: no | | Random Number Generation (RNG) task, assessing inhibition and working memory updating (N =460)

Cognitive Assessment System, assessing attention (N =460) | 50% | Moderate | PA:
intervention = control for mean heart rate and % time spent in moderate PA

CP:
- intervention > control, for the cognitive test measuring inhibition (RNG task)
 - intervention = control all other tests |

Pesce et al., 2016b	RCT, Italy Duration: 6 months Type: Intervention Experimental life skills program integrated into a multisport approach to PE Frequency: 2/week Bout duration: 1 hour Intensity: various intensities Control: Traditional PE	Age: mean 14.5 years N = 109 Senior high school, 7 classes (4 experimental/3 control) volunteered for participation	Accounted for physical fitness, ball-sport skill, executive function and life skills	Assessors: yes Children: yes		Inhibition and updating information/ working memory using random number generation	83%	Moderate	PA: not measured CP: intervention > control for inhibition, favouring the intervention = control for all other outcomes measures
Quach et al., 2016	RCT, US Duration: 8 weeks Type: Interventions (replacing regular PE lessons) (1) Meditation (2) Hatha Yoga Daily home practice Frequency: 2 days/week	Age: mean 13.2 years (12-15) Public junior high school; N = 198	Accounted for age, grade, class period, race, yoga or mediation experience and number of extracurricular activities	Assessors: can't tell Children: can't tell		Working Memory using Automated Operation Span Task (OSPAN score)	87%	Moderate	PA: logbook to monitor home practice CP: intervention = control

	Bout duration: 45 min Control Regular PE lessons								
Telles et al., 2013	RCT, Yoga and physical exercise intervention, India Duration: 3 months Type: yoga or physical exercise Bout duration: 45 min Frequency: 5 days/week Intensity: can't tell	Age: 8-13 years N = 98 ~ 39% female	Accounted for gender, age, BMI and SES	Assessors: yes Participants: can't tell	Teacher's rating of AP	Stroop Color-naming task	>85%	Moderate	PA: no information AP: yoga = physical exercise CP: yoga = physical exercise
Alesi et al., 2016	Pre-post design, After-school Football Exercise Program Intervention + Control, Italy Duration: 6 months Type: football exercise program	Age: 8-10 N = 44 100% male	Accounted for age, SES and baseline CP	Assessors: can't tell Participants: can't tell		BVN 5-11 test battery, including the Forward and Backward Digit Span, the Corsi Block Test, the Visual Discriminatio	Can't tell	Weak	PA: intervention > control for agility skills. CP: intervention > control for posttest scores at the Corsi Block Tapping Task and

	Bout duration: 75 min Frequency: 2 days/week Intensity: MVPA Control: standard 2 x PE lesson per week					n task and the Tower of London			Visual Discrimination accuracy and time intervention = control for posttest scores at the Backward Digit Span and Tower of London task
Ardoy et al., 2014	CT, Education for Fitness (EDUFIT) intervention, Spain Duration: 4 months Type: group 1: PE, group 2: intensive PE Bout duration: 55 min Frequency: 4 days/week Intensity: MVPA Control: 2 x 55 min PE lessons per week	Age: 12-14 years N = 67 ~ 36% female	Accounted for gender, age, pubertal timing, attendance and baseline performance	Assessors: no Participants: can't tell	Average score of core school subjects (e.g. mathematics, language and other subjects such as English, natural sciences)	The medium version of the Spanish Overall and Factorial Intelligence Test	Cohort retention: ~92% Analytical sample: >70%	Weak	PA: Group 2 > Group 1 and control group AP/CP: Group 2 > Group 1 and control group, for indicators of cognitive performance and overall academic achievement.

Coe et al., 2006	CT, Sports, Intervention, US Duration: Approximately 4 months Type: PE Bout duration: 55 min Frequency: 5 days/week Intensity: MVPA Control: no PE during 4 months	Age: 11-12 years N = 229	Accounted for age, gender and ethnicity	Assessors: can't tell Participants: can't tell	Unstandardized grades (mathematics, science, English and world studies) Standardized academic achievement scores (Terra Nova; language, mathematics, science and social studies)		>90% 49% female	Weak	PA: no difference between groups AP: intervention = control
Fredericks et al., 2006	CT, intervention, South Africa Duration: 8 weeks Type: group 1: balance and coordination (experimental), group 2: free roaming (play), group 3: playing	Age: 5 -7 years N = 58	Accounted for gender, age, native language and baseline AP/CP	Assessors: can't tell Participants: can't tell	Grades in drawing, reading and mathematics	Aptitude test for school beginners (8 subtests measuring various cognitive domains)	>90% 60% female	Weak	PA: no information AP/CP: experimental > play/toy for performance on a spatial test, reading and math than the groups

	with table-top toys while seated (toy) Bout duration: 20 min Frequency: 5 days/week Intensity: can't tell Control: regular PE (C)								
Hsieh et al., 2017	CT, Taiwan Duration: 8 weeks Type: After-school gymnastics lessons (gross motor skills) in the weekends Frequency: 2 days/week Bout duration: 90 min Intensity: MVPA Control: maintain daily routines	Age: 7-10 years N = 44	Accounted for age, sex, BMI, SES, IQ, amount of PA, physical fitness, motor skills and extracurricular classes	Assessors: Yes Participants: can't tell		1. Spatial working memory using a modified version of the delayed match-to-sample test (accuracy and reaction time) 2. Event-related potential (P3 component): amplitude and latency	Can't tell	Weak	PA: heart rate monitor (136 bpm) in subsample and ratings of perceived exertion (11.8) CP: response accuracy improved regardless following intervention regardless working memory demands. P3 amplitude was larger at the parietal site after gymnastics training regardless of the task difficulty

Hollar et al., 2010	Quasi-experimental design, Healthier Options for Public Schoolchildren (HOPS) Intervention, US Duration: 2 years Type: classroom-based PA Bout duration: 10-15 min Frequency: 5 days/week Intensity: can't tell Control: 1 school of similar SES and demographics to the intervention schools	Age: 6-13 years N = 4588	Accounted for gender, age and baseline AP	Assessors: can't tell Participants: can't tell	Florida Comprehensive Achievement Test		<30% [30% refers to total analytical sample, AP measured in a subsample of 350 children] % female (can't tell)	Weak	PA: not objectively measured and introduced in the second year as part of the intervention AP: intervention > control for mathematic scores
Hunter et al., 2014	RCT, Active Kids Active Minds (AKAM), Intervention, Australia Duration: 5 months Type: aerobic (PE, Smart Moves (PA	Age: 9-10 N = 107 50% female	Accounted for gender, age, SES (same school) and baseline CP	Assessors: yes Participants: can't tell	Average school grades for the first two semesters	Cognitive assessment System (CAS) for planning, attention, simultaneous and successive processes	60%	Weak	PA: No direct comparison between the groups. PA at posttest is lower than at pretest, for both groups. The intervention group had lower levels of PA than the control.

	delivered by classroom teacher) and AKAM (running, games, etc.) Bout duration: 60 min Frequency: 5 days/week Intensity: moderate Control: Approximately 5x30 min a school week of PA through PE and Smart moves								*Mean time spent in moderate intensity in the intervention group during AKAM classes was 15 min/day.* AP: intervention = control CP: intervention = control
Lakes et al., 2013	CT, Healthy for Life Intervention, US Duration: 9 months Type: PE and Taekwondo Bout duration: 40-45 min Frequency: 5 days/week (3 PE and 2 Taekwondo) Intensity: moderate Control:	Age: 12-13 years N = 60 ~50% female	Accounted for gender, age, BMI, ethnicity and baseline questionnaire scores	Assessors: yes Participants: can't tell		Hearts and Flowers Executive Function Test Parent-Rated Inhibitory (Attentional and Behavioral) Control Questionnair	Can't tell	Weak	PA: no information CP: intervention > control for accuracy (log-transformed) on the congruent trials and for behavioral control (as rated from their parents)

	5 x 40-45 min PE per week					e, known as SWAN)			
Mavilidi et al., 2015	CT, "Embodying knowledge through whole or partial physical exercises" Intervention, Australia Duration: 10 weeks Type: 4 conditions: integrated physical exercise, non-integrated physical exercise, gesturing physical exercise and conventional condition Bout duration: 15 min Frequency: 2 days/week Intensity: 4 conditions: MVPA (partial), MVPA (partial), light, sedentary	Age: 4-5 years N = 125 ~ 49% female	Can't tell	Assessors: can't tell Participants: can't tell		Free-recall and cued-recall test of 14 Italian words	>85%	Weak	PA: Integrated and non-integrated condition > gesturing and conventional condition; for time spent in MVPA CP: Free recall - integrated condition > non-integrated, gesturing and conventional condition Cued recall - integrated condition > gesturing and conventional condition. - non-integrated condition > conventional condition - gesturing condition > conventional condition

Mohammad et al., 2013	Pre-post design, Rope skipping Intervention, Iran Duration: 3 months Type: Rope skipping Bout duration: 45 min Frequency: 1 day/week Intensity: moderate Control: none	Age: 9-10 years N = 84 100% female	Not Applicable	Assessors: can't tell Participants: can't tell	School grades (class point average)		Can't tell	Weak	PA: no information AP: improved school grades after rope skipping
Mullender-Wijnsma et al., 2016	Cluster RCT, Fit and Academically Proficient at school Intervention + Control, Netherlands Duration: 2 years [22 weeks/year] Type: aerobic Bout duration: 20-30 min Frequency: 3 days/week Intensity: MVPA (not measured)	Age: 8-9 N = 499 45% female	Accounted for gender, age, SES (proxy measure) and baseline AP	Assessors: no Participants: no	Two math tests (speed and general math skills), two language tests (reading and spelling)		>70%	Weak	PA: no information AP: - intervention = control for reading after the 1st and 2nd year - intervention = control for spelling and mathematics after the 1st year - intervention > control for spelling and mathematics after the 2nd year

	Control: Regular scholastic curriculum							- intervention > control for ability on math retrieved from a child academic monitoring system after the 1st and the 2nd year
Reed et al., 2013	Pre-posttest design, intervention, US Duration: 9 months Type: aerobic Bout duration: 45 min	Age: 7-14 years N = 470 50% female	Accounted for gender, age and baseline CP	Assessors: no Participants: can't tell		Raven's Standard Progressive Matrices (SPM)	~ 90%	Weak	PA: intervention > control for fitness CP: - Elementary school: Males, but not

	Frequency: 5 days/week Intensity: can't tell Control: Grade 2 to 5: 1 x 45 min/week of PE and grade 6 to 8: 5 x 50 min/ week of PE in the fall semester only					Perceptual speed test			females, improved their scores in certain sections of the Fluid intelligence test (section D) after the intervention as compared to males in the control school. Females, but not males, in the intervention group improved significantly on section 2 and 3 of the perceptual speed test. - Middle school: females in the intervention group improved significantly on some sections of the fluid intelligence test (B, C, D and E). Males in the intervention group showed this effect only for section E
Riley et al., 2016	Cluster RCT, "Encouraging Activity to Stimulate Young	Age: 10-12 years N = 240	Accounted for gender, age, BMI, baseline CP and AP	Assessors: yes	Standardized mathematics	Time spent 'on-task' during the	100%	Weak	PA: intervention > control for counts per minute and time spent in

	(EASY) Minds" Intervention, Australia Duration: 6 weeks Type: Mathematics integrated with PA lessons Bout duration: 60 min Frequency: 3 days/week Intensity: mostly sedentary, light to MVPA Control: Regular scholastic curriculum	41% female		Participants: can't tell	progressive achievement test	mathematics lessons				MVPA (increase of 2.6%). AP: - intervention > control for time spent "on-task" - intervention = control for mathematics aptitude and performance
Sallis et al., 1999	RCT, Sports, Play and Active Recreation for Kids (SPARK) Intervention + Control, US Duration: 2 years Type: PE classes delivered by trained teachers (T) or PE	Age: 9-10 years N = 1538	Accounted for gender, ethnicity and baseline grades	Assessors: yes Participants: can't tell	Standardized academic achievement scores (MAT6 and MAT7)		<50% ~48% female	Weak	PA: no information AP: All scores declined from baseline to posttest, except the reading scores in the S group in the 1st cohort, which increased.

	specialist (S); 15 min of health-fitness and 15 min of skill-fitness activity Bout duration: 30 min Frequency: 3 days/week Intensity: can't tell Control: regular PE (C)								-Basic battery: T < S, C (2nd cohort only) - Language: T < C (1st cohort), T, C < S (2nd cohort) - Mathematics: no differences - Reading: S > C (1st cohort), T < C (2nd cohort)
Shore et al., 2014	CT, School-based pedometer intervention (SBPI) referred to as the Helping Youth Pursue Physical Activity and Exercise (HYPPE) Intervention, US Duration: 6 weeks Type: Increased PA opportunities, including access to group lessons (martial arts, sport specific training),	Age: 11-12 years N = 113 ~58% female	Accounted for gender, age, ethnicity and baseline accrued steps	Assessors: can't tell Participants: can't tell	Unstandardized grade point average (GPA)		>80%	Weak	PA: intervention > control for number of accrued daily steps (12307 vs 10608). Furthermore, there were no intervention related improvements in physical fitness AP: intervention = control for academic achievement

	daily 'prompts' to achieve a specific amount of steps, enhanced PE curriculum (10 lessons)and lifestyle training Bout duration: achieve 10.000 steps/day Frequency: 5 days/week Intensity: can't tell Control: Standard school curriculum, plus encouragement to achieve 2000 steps a day								
Sjowall et al., 2017	CT, Sweden Duration: 2 school years Type: Besides regular PE (2 x 60 min) additional high intensity PA	Age: 6-13 years 2 schools N = 470	Accounted for gender, age, fitness and baseline values CP	Assessors: can't tell Participants: can't tell		Working Memory test using (spatial and verbal), arithmetic task	Can't tell	Weak	information = control

	Frequency: 3 days/week Bout duration: 40 min Intensity: high intensity Control: Regular PE (2 x 60 min)								
Spitzer and Hollmann, 2013	Quasi-experimental design, intervention, Germany Duration: 4 months Type: exercise (e.g. basketball, soccer, hockey, handball) and dance lessons (hip-hop) Bout duration: 30 min Frequency: 3 days/week Intensity: can't tell Control: regular PE lessons	Age (mean) 12.5 (intervention) and 13 (control) years [range unknown] N = 44	Can't tell	Assessors: no for school grades, can't tell for the d2 test Participants: can't tell	School grades in German (literacy), English and mathematics	D2 test of attention	Can't tell	Weak	PA: no information AP: intervention > control for German grades. No intervention effect for mathematics or English grades. CP: intervention = control

	Quasi-experimental design, intervention, Germany Duration: 4 months Type: games Bout duration: 30 min Frequency: 3 days/week Intensity: can't tell Control: regular PE lessons	Age: mean 12.4 (intervention) and 12.3 (control) years [range unknown] N = 148	Can't tell	Assessors: no for school grades, can't tell for the d2 test Participants: can't tell	School grades in German (literacy), English and mathematics	D2 test of attention	<60% ~38% female	Weak	PA: no information AP: intervention > control for mathematics grades; intervention = control for English and German grades; CP: intervention > control for d2 scores
	Pre-posttest design, intervention, Germany Duration: 4 months Type: aerobic Bout duration: can't tell Frequency: 1 day/week Intensity: MVPA Control: none	Age: mean 11.6 years [range unknown] N = 53 40% female	Can't tell	Assessors: no for school grades, can't tell for the d2 test Participants: can't tell		Teacher's observation on focus and concentration	Can't tell	Weak	PA: no information CP: Teachers reported that children were more focused after jogging

| Tarp et al., 2016 | Cluster-RCT, Learning, Cognition and Motion (LoCoMotion), Denmark

Duration: 20 weeks
Different components
1) PA in academic subjects (daily)
2) scheduled PA during recess (weekly)
3) PA homework (5-10 min) (daily)
4) active transportation (2-week cycling campaign) (daily)

Intensity: various intensities

Control: normal practices (no additional PA) | Age: mean 12.7-13.1 years

16 schools throughout Denmark (13 schools recruited by external collaborator and 3 schools who volunteered)

N = 855 | Accounted for ethnic origin, pubertal development, baseline AP and CP and overweight status | Assessors: no

Participants: no | Custom made mathematics test | Cognitive control measured by a modified Eriksen flanker task | >65% | Weak | PA: intervention = control for overall PA levels and daily min spent in MVPA

AP: intervention = control for mathematics skills

CP:
- intervention = control for response accuracy and reaction time on congruent or incongruent trials
- intervention = control for accuracy interference; control > intervention for reaction time interference scores. (larger reduction in interference score) |

RCT = randomized controlled trial; CT = controlled trial [both RCT or CT designs are classified as strong]; MVPA=moderate to vigorous physical activity, PE=physical education, PA=physical activity; Sample refers to the characteristics of the target population at baseline, * = undecided due to study sample being part of larger dataset reported elsewhere; AP=academic performance; CP=cognitive performance (categories chosen based on information in article); GPA = Grade Point Average; Confounders refer to possible differences in baseline between the intervention and

control group, if and how authors have accounted for differences (e.g. baseline AP/CP, SES, age, gender); Blinding: yes= assessor/participants are not aware of the research question, no=assessors/participants are aware of the research question, can't tell = there is no information on whether the assessors/participants were blinded (in case (A) assessors and children are not blinded, or (B) assessors are not blinded and information on children is missing (can't tell) this item is scored 'weak'; in case both children and assessors are rated 'can't tell' this item was scored 'moderate'; in case assessors were blinded this item was scored as 'strong'); Analytical sample refers to the percentage of the sample that was included in the analysis, Cohort retention refers to the % of participants who completed the study;

Quality refers to the global rating of the intervention (STRONG (at least two high-quality ratings, no weak ratings), MODERATE (less than two strong ratings, no more than one weak rating) and WEAK (two or more weak ratings);

Colors in table refer to the quality rating: green = strong; blue = moderate; red = weak