Effects of physical activity interventions on cognitive and academic performance in children and adolescents: a novel combination of a systematic review and recommendations from an expert panel

Amika S Singh,1 Emi Saliasi,1 Vera van den Berg,1 Léonie Uijtdewilligen,2 Renate H M de Groot,3 Jelle Jolles,4 Lars B Andersen,5 Richard Bailey,6 Yu-Kai Chang,7 Adele Diamond,8 Ingegerd Ericsson,9 Jennifer L Etnier,10 Alicia L Fedewa,11 Charles H Hillman,12 Terry McMorris,13 Caterina Pesce,14 Uwe Pühse,15 Phillip D Tomporowski,16 Mai J M Chinapaw1

ABSTRACT

Objective To summarise the current evidence on the effects of physical activity (PA) interventions on cognitive and academic performance in children, and formulate research priorities and recommendations.

Design Systematic review (following PRISMA guidelines) with a methodological quality assessment and an international expert panel. We based the evaluation of the consistency of the scientific evidence on the findings reported in studies rated as of high methodological quality.

Data sources PubMed, PsycINFO, Cochrane Central, Web of Science, ERIC, and SPORTDiscus.

Eligibility criteria for selecting studies PA-intervention studies in children with at least one cognitive or academic performance assessment.

Results Eleven (19%) of 58 included intervention studies received a high-quality rating for methodological quality: four assessed effects of PA interventions on cognitive performance, six assessed effects on academic performance, and one on both. All high-quality studies contrasted the effects of additional/adapted PA activities with regular curriculum activities. For cognitive performance 10 of 21 (48%) constructs analysed showed statistically significant beneficial intervention effects of PA, while for academic performance, 15 of 25 (60%) analyses found a significant beneficial effect of PA. Across all five studies assessing PA effects on cognitive or academic performance assessment, measurement of compliance and academic performance in children, and formulate research priorities and recommendations.

The expert panel confirmed that more ‘high-quality’ results are needed to elucidate the nature of the relationship. Several hypotheses have been proposed to explain posited PA-related effects on cognitive and academic performance. Regular PA alters neurogenesis and angiogenesis and enhances central nervous system metabolism. Furthermore, it

INTRODUCTION

Cognitive skills are crucial for school readiness and academic performance,1–3 and are supported by a variety of brain regions, which continue to mature throughout adolescence.4 This developmental window poses a great opportunity for experience-dependent plasticity,5 as the structural and functional organisation of the brain can be positively influenced through enriched environmental conditions6 like, for example, physical activity (PA). As such, it is critical to advance our understanding of the potential to positively influence brain development.

The physical and mental health benefits of PA are widely acknowledged,7 8 but less is known about the potential effects of PA on cognitive and academic performance. A number of reviews and reports on this topic have been published during the last decade,9–19 mainly concluding that PA is positively associated with cognition and with structural and functional brain health and a neutral association with academic performance for children.12 Not all of these reviews were systematic, took into account the methodological quality of the studies included, and were exclusively looking at intervention studies. Above all, most reviews summarised the existing evidence at study level. So, if a study reported a significant association of PA with one, but not all, outcome measures, it was categorised as a ‘positive’ study. In addition, most reviews concluded that many aspects of this relationship remain unclear and that further research is needed to elucidate the nature of the relationship.

Several hypotheses have been proposed to explain posited PA-related effects on cognitive and academic performance. Regular PA alters neurogenesis and angiogenesis and enhances central nervous system metabolism.20 Furthermore, it
has been suggested that regular PA increases the availability of certain growth factors, such as brain-derived neurotrophic factor (BDNF), insulin-like growth factor-1 (IGF-1) and vascular endothelial growth factor (VEGF). All these systems and factors are involved with the maintenance and plasticity of the structure and function of the brain, and thus with learning and memory.

The cognitive and academic training of children is, in large part, a task entrusted to the educational system. Within the school context, the academic performance of pupils is typically rated through the assessment of their knowledge and scholastic aptitude in various subjects, the most prominent of which are mathematics and literacy. To enhance academic performance, instructional time for core academic subjects is prolonged and protected, often at the expense of time spent in physical education (PE) and other areas of the curriculum. PE has been attributed a lower status than academic subjects and is perceived by some to interfere with academic performance. There is, however, no evidence indicating that increased time spent in PA in the school setting has an adverse effect on academic performance.

To foster progression in this research field we: (1) systematically summarised current experimental evidence on the effects of PA interventions on cognitive and academic performance in children, considering the methodological quality of the studies; (2) conducted a Delphi study among acknowledged researchers in this field to identify gaps in knowledge, compile a list of future research priorities, and generate recommendations for future research.

METHODS

The current review was registered at PROSPERO (CRD42017082303). We applied the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) reporting guidelines for this review.

This study followed a two-step procedure, with both steps taking place in parallel: (1) a systematic literature review of intervention studies investigating the effects of regular PA on cognitive and/or academic performance in children and adolescents (0–18 years); (2) a Delphi study among scientific experts on the topic of PA and cognitive/academic performance.

Systematic review

Literature search and eligibility

Randomised and non-randomised intervention studies that investigated the effects of PA on cognitive/academic performance in children were identified from a systematic search of the literature using six electronic bibliographic databases (PubMed, PsycINFO, Cochrane Central, Web of Science, ERIC, and SPORTDiscus), until September 2017. Search terms were related to the major topic and free text words in the title, and were adapted to each specific database. Information on the search strategy is included as an appendix to this manuscript (see online supplementary appendix 1). Studies were considered eligible if they met the following criteria: (1) they were PA-related intervention studies; (2) they had at least one cognitive or academic performance assessment; (3) the population sample consisted of apparently healthy children or adolescents. We included only full-text articles published in English-language peer-reviewed journals. For purposes of generalizability, studies that focused on a clinical sample (eg, overweight/obese children, Attention-Deficit Hyperactivity Disorder (ADHD)) were excluded.

Data extraction

One researcher (ES, VB, or AS) selected studies of potential relevance based on titles, abstracts and inclusion criteria. Reasons for studies to be excluded in this phase were mainly that it was a clinical sample, a cross-sectional study design or it assessed acute effects of PA. The researchers in the expert panels were asked to cross-check the retrieved studies to ensure that no important studies were missing. Thereafter, two reviewers independently checked whether the full text studies met the inclusion criteria and extracted data (either ES and VB or AS and LU, RdG, JJ, MC, or VB). A total of 58 intervention studies fulfilled the inclusion criteria.

Methodological quality assessment and synthesis of scientific evidence

The methodological quality of the included studies was rated using an adapted version of the ‘Quality assessment tool for quantitative studies’, that assesses study quality based on predefined criteria. After relevant studies were identified, four authors (either ES and VB or AS and MC) independently assessed the methodological quality in seven dimensions—that is, selection bias, study design, adjustment for potential confounders (eg, age, gender, social economic status, and baseline academic/cognitive performance), data collection methods, blinding, withdrawals, and dropouts. Discrepancies between authors were discussed and resolved by consensus. In case of disagreement, a third reviewer who was not involved in the scoring was consulted (AS, MC or VB).

The assessment tool used for the methodological assessment of the intervention studies defines high-quality studies as having at least two strong and no weak dimensions, moderate-quality studies as having less than two strong dimensions, but no more than one weak dimension, and low-quality studies as having more than one weak dimension.

In accordance with a previous systematic review by our group, the level of scientific evidence was rated based on the following criteria:

- Strong evidence, provided by a minimum of two high-quality (ie, strong rating) studies with generally consistent findings
- Moderate evidence, provided by a minimum of one high-quality study and one or more ‘lower quality’ (ie, moderate or weak rating) studies with generally consistent findings
- Inconclusive evidence, when only one study was available or when findings were not consistent in two or more studies.

Findings were considered to be consistent when at least 75% of the studies reported statistically significant results in the same direction. When two or more studies of high methodological quality were identified, the studies of moderate or weak quality were not considered in the evaluation of the consistency of the scientific evidence. Evidence synthesis took place at the level of construct for cognitive performance (eg, attention inhibition) and outcome for academic performance (eg, mathematics).

Expert panel

Structured communication (Delphi) method

The leading authors (AS, ES, VB, MC) independently recommended a multidisciplinary group of international experts on the topic of PA and cognitive/academic performance in children from...
different regions of the world. This initial sample of experts was asked to recommend additional experts. Experts were assigned to two separate panels (panel A: senior position; panel B: junior or mid career position or previously affiliated with research groups of experts in panel A).

Procedure
The opinions of the experts on the topic of the relationship between PA and cognitive/academic performance in children were systematically gathered through a modified Delphi technique, similar to the procedure of Gillis and colleagues. In the first round, we asked the experts in panel A to provide a minimum of five alternatives in response to the following question: “In your opinion, what are the most important theoretical and methodological questions that need to be addressed in order to elucidate the relationship between physical activity and cognitive and academic performance in children and adolescents?” Each expert in panel A received a link to a digital survey via email and was asked to provide five research questions. Two authors (ES, AS) collected and summarised the suggestions of panel A. Reoccurring or similar research questions were combined into one. In the second round, experts from both panels (A and B) received a link to a digital questionnaire containing the summarised responses from round 1 and were asked to rate them on importance for generating new scientific evidence on a 5-point Likert scale (5=very important, 4=important, 3=moderately important, 2=of little importance and 1=unimportant). The cumulative score for each item was calculated. Next, the issues were organised into clusters, according to the overarching topics they addressed. This information was presented to the experts in the final round, wherein they received the clustered research questions and were asked whether they agreed with the proposed clusters. Experts were blinded to the individual answers of other experts in the panels.

RESULTS
Systematic review
From the 8103 identified references from the six electronic databases, 58 studies met the inclusion criteria (figure 1).

General study characteristics
The general description of the studies, their methodological characteristics, rating and main results are summarised in the online supplementary table 1.

Of the identified intervention studies, 20 were conducted in the USA and 23 in Europe (Italy (n=6), Sweden (n=4), Netherlands (n=3), Denmark (n=2), Germany (n=2), Norway (n=2), UK (n=2), Spain (n=1), and Switzerland (n=1)). The other studies were conducted in Australia (n=6), India (n=3), Taiwan (n=2), Canada (n=1), South Africa (n=1), Iran (n=1) and New Zealand (n=1).

The number of participants included in the intervention studies ranged from 16 to 1953, and participants’ ages at

---

**Figure 1** Flow of literature selection.
baseline ranged from 3.51 to 16 years. The majority of the studies included in the current review assessed the effects of PA in children aged 6–12 years. Seven studies reported children aged 5 years or younger, and 13 studies included youths (ie, older than 12 years) in their studies. Most intervention studies had a strong experimental design, with 27 studies using an RCT design or cluster RCT design, and 21 studies using a controlled trial (CT) design. Five studies used a quasi-experimental design, four a pre-post-test, and one a cohort analytic design. Two studies, reporting on the effects of three substudies each, applied different designs per substudy.

The shortest intervention duration was 1 week, and the longest spanned a period of 9 years. The lowest PA dose delivered was 10 min per week. The FITKids intervention delivered the highest dose of PA (ie, 350 min per week in exercise sessions, made up of 70 min for 5 days a week). The majority of the intervention studies implemented PA sessions of at least 30 min (n=41). Seventeen interventions assessed the effects of PA sessions with a shorter duration, varying from 5 to 30 consecutive minutes. Most interventions delivered PA in one session per day, except for seven studies in which either two sessions per day were delivered, or PA was spread throughout the day aiming to reach at least 10,000 steps. Many studies applied interventions that were delivered 5 days a week (n=28). The lowest frequency was 13 weeks, and the highest frequency was 6 days per week. Most studies (n=34) assessed the effects of PA consisting primarily of aerobic exercise. Five studies contrasted the effects of yoga to aerobic exercise, and nine studies implemented more cognitively engaging PA, consisting of either coordinate exercise, skill-based training, gesturing physical exercise, perceptual motor training, and motor-enriched learning activities. Six studies assessed the effects of motor demanding or cognitively engaging activities. Three of these studies implemented PA games specifically tailored to challenge core executive functions in a gross-motor fashion. The majority of studies implemented separate PA sessions, whereas 14 studies delivered physically active academic lessons.

**Measurements of cognitive and academic performance**

Eleven studies reported on unstandardised school grades on subjects like mathematics and language as measure of academic performance. The other 19 studies that reported effects on academic performance assessed effects on national curriculum levels, standardised school grades (eg, grade point average), or standardised performance tests (eg, Wechsler Individual Achievement Test, Canadian Achievement Test). A small number of studies assessed academic performance on other subjects, such as drawing or English as a foreign language. Most studies reporting intervention effects on cognitive performance, assessed domains of executive functions (in particular inhibition, working memory, updating, attention, task switching, and planning), speed of information processing, fluid and crystallised intelligence.

**Methodological quality**

Eleven studies were rated as of high methodological quality, and 17 as of weak methodological quality (see online supplementary table 1). McClelland et al reported on three substudies, resulting in categorisation of one study as moderate/weak, which reflects the differences in quality among the three substudies. For the evidence synthesis, only studies of high methodological quality were taken into account.

**Evidence synthesis**

Of the 11 high-quality studies, four reported outcomes on cognitive performance, and one on both. We present the outcomes of the 11 high-quality studies in the online supplementary table 2.

Ten out of 21 (48%) analyses in five high-quality studies examining the effect of PA on cognitive performance found a significant beneficial intervention effect, resulting in inconclusive evidence. For intervention effects of PA on academic performance, 15 out of 25 (60%) analyses constructed in six high-quality studies found a significant beneficial effect, also leading to inconclusive evidence. Stratifying the evidence synthesis at the level of outcome for academic performance we conclude that there is strong evidence for beneficial effects of PA on maths performance (beneficial effects on 86% of the outcomes), but inconclusive evidence for language performance (beneficial effect on 27% of the outcomes).

None of the studies reported significant adverse effects of PA on cognitive and academic performance. The studies are now briefly described based on the extent to which the findings were supportive of beneficial effects of exercise. Donnelly and colleagues found improved academic performance on all reported outcome measures (ie, the composite scores of reading, spelling and mathematics), when comparing children who received daily bouts of moderate to vigorous PA during academic lessons for 3 years (up to 90 min/week) with children who followed the regular curriculum. Other studies found selective effects of PA interventions on specific measures of cognitive and academic performance. For example, Ericsson found higher grades in literacy and mathematics in children who participated in a 45 min PE session each school day and an optional modified motor training of 60 min per week for 3 years, than in children following the regular PE curriculum at 1 and 2 year follow-up. Intervention effects on literacy were not sustained into the third year. The intervention also showed small significant beneficial effects on attentional performance and impulse control after 1 and 2 years, which were not sustained into the third year. After 9 years, children in the intervention group had significantly higher qualification scores for upper school.

Participation in an exergaming-based intervention (2×15 min, three times a week) in a study by Gao et al resulted in higher maths grades in fourth-grade children compared with a group receiving unstructured recess, but no significant differences in reading scores. Kirk et al found that preschoolers who participated in twice-daily 15 min physically active academic lessons during 6 months performed significantly better on picture naming and alliteration tasks than the group who participated in regular academic lessons. In contrast, preschoolers in the control and intervention groups showed similar performance on the rhyming task. Telford et al compared 2 years of specialist-taught PE with non-specialist taught PE and found a significant intervention effect of 10.9 points in numeracy but not in reading or writing in 8- to 11-year-old children.

Hillman and colleagues reported more improvement from pre-test to post-test in some aspects of cognitive performance and measures of brain function in children aged 8 to 9 years who followed the FITKids intervention versus a control group. In particular, enhanced performance accuracy and increased P3 amplitude were observed in tasks requiring higher amounts of attention, inhibition and cognitive flexibility. Tasks requiring less executive control were unaffected by the FITKids intervention.
Schmidt et al\textsuperscript{54} found that in 10- to 12-year-old children the group that participated in a cognitively- and physically-demanding exercise intervention (team games) showed more improvement from pre-test to post-test on cognitive flexibility (ie, being able to shift from one task to another) than the group receiving an aerobic intervention with low cognitive demands or the group receiving standard PE with low physical and low cognitive demands. Performance on the other two executive function subcomponents (ie, updating and inhibition) were similar between groups. Subramaniam et al\textsuperscript{55} reported beneficial effects of structured PA versus unstructured physical and creative sedentary activities in 12- to 17-year-olds on tasks measuring attention, concentration, non-verbal fluency, and mental flexibility. Both groups received a comparable dose (ie, 120 min, 6 days per week) during the 6 month intervention period and showed improved performance compared with baseline, with significantly better scores of the group receiving structured PA on all constructs measured.

Three high-quality studies found no significant intervention effects on any of the assessed outcome measures. Ahamed et al\textsuperscript{56} found no significant intervention effect on academic performance in 9- to 11-year-old children. This intervention included 15 min additional classroom-based PA 5 days a week for 16 months. Chaya and colleagues\textsuperscript{36} compared daily yoga with structured PA (stretching and aerobic exercise) and found no significant differences on any of the cognitive performance tests between groups. Resaland et al\textsuperscript{54} found no significant intervention effect on mathematics, reading or English in 10-year-old children. Analyses of the subgroup who performed worse at baseline for numeracy showed significant beneficial effects of the intervention. The 7-month intervention consisted of physically active lessons, active breaks between lessons, and active homework.

Five high-quality studies examined the effects on academic performance measures indicators related to mathematics.\textsuperscript{38,41,42,70,83} Across all five studies, beneficial effects were reported in six out of seven (86\%) outcomes. Six high-quality studies that examined effects on academic performance, measured indicators related to language (eg, literacy/reading/rhyming).\textsuperscript{38,41,42,51,70,83} Three of these studies reported significant beneficial effects in eight out of all 15 (53\%) outcomes assessed.

All high-quality studies assessed the effects of an adapted curriculum (eg, increased frequency or duration of the PE session, adapted content of PE lessons, structured recess) or active academic lessons compared with regular curriculum activities (eg, regular academic and PE lessons, unstructured recess). Most of the high-quality studies used control groups that followed regular PE lessons. Hillman et al\textsuperscript{50} used a wait-list group, Gao et al\textsuperscript{58} and Subramaniam et al\textsuperscript{55} compared structured versus unstructured recess, and Kirk et al\textsuperscript{51} compared physically activity lessons in pre-schoolers to lessons without an activity component.

**Expert panel**

**Expert demographics**

The initial sample of experts suggested by the leading authors of this manuscript, provided names of 33 other experts. Experts were assigned to two separate panels: 16 researchers with a senior position were assigned to panel A; the 17 researchers who were assigned to panel B held a junior or mid career position or had been previously affiliated with research groups of experts in panel A. Of the experts invited to join panel A, 12 agreed to participate, three declined and one expert did not respond to the email and reminder. For panel B, 11 researchers accepted the invitation, one declined, and five did not respond. See the online supplementary appendix 2 for the experts participating in panels A and B.

The 23 experts participating in our panels were affiliated with educational organisations located in Asia, Australia, Europe and North America (figure 2), and were in different stages of their academic careers, ranging from the last stages of their PhD to full professors. Experts in panel A had an average of 23 years of research experience, of which 17 years were dedicated to investigating the association between PA and cognitive and academic performance in children. Experts in panel B had an average of 9 years of research experience with 7 years on PA and cognitive and/or academic performance in children.

**Expert panel results**

Twelve experts (panel A) completed the first round; 11 experts nominated five research questions each and one expert nominated four, resulting in a total of 59 questions. After removing duplicating research questions, for the second round, we put forward 46 questions to panels A and B to be rated. All the experts agreed with the final formulation and clustering of the research questions (see online supplementary table 3).

The highest rated research question (111 out of 115 points) highlighted the need to establish the causality of the relationship between PA and cognitive and academic performance. This research question was grouped in cluster 1 ‘Causality’ (total number of research questions in this cluster n=8). The remaining research questions were grouped into four clusters, addressing the following overarching topics: PA characteristics (n=18), moderators (n=8), and mechanisms (n=7). Seven research questions did not fit into one of the above mentioned clusters, therefore we formed a ‘miscellaneous’ cluster. Two questions fitted two clusters, and were included in both.

**DISCUSSION**

We summarised the current evidence on the effects of PA interventions on cognitive and academic performance in children by conducting a systematic literature review including a methodological quality assessment. Hereafter, we discuss the review’s main findings in light of the recommendations of our expert panel.

Six high-quality studies included in our systematic review examined the effect of PA on cognitive performance constructs: 48\% of the analyses showed a significant beneficial effect of PA, resulting in inconclusive evidence for beneficial PA effects on cognitive performance. Fifteen out of 25 (60\%) analysed academic performance outcomes in seven high-quality PA intervention studies were significant, also leading to inconclusive evidence at a summary level.

When stratifying at the level of outcome for academic performance, we found strong evidence for beneficial effects of PA on maths performance. There is inconclusive evidence for beneficial effects of PA on tests assessing indicators related to language.

It is important to keep in mind that many studies compared various types and doses of physical activity against one another and did not include true no-PA control conditions. This resulted in a large variation in intervention contents, as well as contrasts between intervention and control groups. As such, our conclusion refers to the effects of additional PA or adaptations of the PA curriculum such as structured recess, frequency/intensity/specialist taught PE, active academic lessons on cognitive or academic performance, but not on the effects of PA per se.
Although we found no evidence for beneficial effects on cognitive performance or overall academic performance, we found strong evidence for beneficial effects of PA on maths performance. This finding is in line with Fedewa’s and Ahn’s meta-analysis, in which they conclude that the largest effectiveness of PA was found for maths performance. Interestingly, all four studies that reported beneficial effects on maths performance assessed intervention programmes with a minimal frequency of three sessions per week and a minimal programme duration of two school years.

This inconsistency in our findings is echoed by the call of our expert panel for more well-designed studies that provide insight into the causality, mechanisms and moderators of this effect, as well as the PA characteristics that optimally benefit cognitive/academic performance. More insight into the causal relationship and the effects of different parameters of PA (e.g., duration, frequency, sort of PA, timing of PA) could be established by high-quality RCTs with multiple PA intervention arms as well as a non-active control group.

To establish the effect of combining academic content and PA, there should be at least three groups: (a) the intervention group that combines PA with academic content; (b) the non-active control group that only receives the same academic content as the intervention group, but without PA; (c) the active control group that only receives the PA component. Up to now, many studies with a programme combining academic content with PA compared the intervention group to a regular curriculum control group with a different academic content, which does not allow us to draw clear conclusions on the added value of combining PA with academic content as the academic content also differs between groups. The aforementioned design would allow us to test (1) for causality because of random assignment and the manipulation of the independent variables and (2) for main effects and interactions simultaneously.

Cognitive/academic performance can theoretically improve through biological or psychosocial pathways, and an interplay between the two is possible and probable. Evidence to support these proposed mechanisms is sparse. Our expert panel highlighted the need to understand PA effects not only at the behavioural level, but also at the cellular, functional and morphological level. A small number of studies examined functional brain correlates that may underlie the effects of PA on cognitive performance. Expanding on these findings, Chaddock and colleagues stated that a PA intervention may result in a more adult-like recruitment of prefrontal brain regions, which is important for many aspects of executive functions. As confirmed by our expert panel, more research is needed to understand the mechanisms underlying the potential effects of PA on cognitive performance. Specific suggestions related to elucidating PA-related effects on BDNF and other cellular and molecular mechanisms, exploring the potential mediating role of psychological mediators and advancing our understanding of the potential roles of social belonging and support in improving cognitive and academic performance.

Our expert panel also emphasised the need to distinguish between the different qualitative and quantitative characteristics of PA, and the differential effects that these characteristics
may have on various aspects of cognitive and academic functioning. For instance, the effect of different types of PA on cognitive performance needs further investigation. One high-quality study suggested larger effects of coordinative or perceptual–motor forms of exercise on cognitive performance, particularly on executive functions, than aerobic exercise. Differential effects between exercise types may stem from various reasons. Variability of practice that is central to coordination and perceptual–motor exercise training, for example, may be an interface between motor and cognitive development promotion. Further research is needed to establish whether more cognitively engaging forms of exercise have larger effects on cognitive and academic performance.

Our understanding of the relationship between PA and cognitive and academic performance can also be further improved by exploring possible mediators and moderators (eg, baseline PA and cognitive performance). The evidence regarding possible mediators, including proposed biological or psychosocial mediators, is scarce. Only one study conducted by Pesce et al conducted a mediation analysis and found that ball skills mediated the effect of an intervention programme during physical education on inhibition, a core executive function.

Regarding moderators, it is likely that some children may benefit more from additional PA such as children with low PA levels or overweight children. For example, Subramaniam et al found larger benefits in non-athletes and Crova et al found PA to be more beneficial to overweight children in the intervention group.

The expert panel also indicated the need to examine the potential effects of limiting or interrupting sedentary behaviour on cognitive and academic performance. Most examined PA interventions require a high level of commitment from school staff to be implemented. Interrupting and limiting sedentary time with standing desks or short activity bouts may be more feasible in the school setting, and, therefore, a promising form of intervention.

Summarising the expert opinion, we conclude that more well-designed studies are needed to: (1) assess the effects of PA on cognitive and academic performance; (2) specify PA characteristics that affect cognitive of academic performance; (3) understand underlying mechanisms and moderators of this effect.

**Strengths and limitations**

This study employed a structured Delphi technique to map the research priorities set and ranked by international experts. An important strength of this method is the anonymity of responses throughout the rounds. Experts were not influenced by one another in generating or rating the responses, and, in fact, were not known to each other until after data collection was completed. These steps were taken to assure greater objectivity and generalisability in rating the research questions. In addition, experts on our panel represented research institutions in various countries and continents and offered experience from different scientific fields.

A major strength of our review is that we only included intervention studies, which enables us to draw conclusions on the effects of PA interventions on cognitive and academic performance. Another strength is the combination of the systematic literature search with a methodological quality assessment: both served as the basis for the evidence synthesis at the level of outcome measures. Previous reviews based their conclusions on study level findings (ie, if a study reported beneficial effects of an intervention on one but not all outcome measures, the study was rated as ‘positive’). We believe that our choice of assessing evidence on the level of outcomes provides a more detailed picture of potential PA effects on cognitive/academic performance.

Two researchers independently rated the methodological quality, and, when necessary, the authors of the included studies were approached to provide additional information to ensure an optimal level of accuracy. While this is a strength of the study, the methodological quality scores of studies included here and in a previous study by Norris and colleagues showed discrepancies in the outcome quality assessment, suggesting a degree of subjectivity. This subjectivity is partly due to the interpretation of the categories, but also to the fact that weighing and interpretation of certain methodological aspects are highly dependent on the specific research question. We have not recorded the percentage agreement between the raters in the different phases (ie, selection of the studies, quality assessment), but all disagreements were resolved by discussion between both raters or if necessary by a third author. Another potential limitation is publication bias: positive findings are more likely to be published, resulting in an overestimation of the beneficial effect of PA on cognitive/academic performance.

**Recommendations**

Less than a quarter of the 58 studies included in our review were rated as of high methodological quality. Based on our analyses, discussions and reflections, we formulated the following recommendations for future research.

a. An important aspect in PA interventions is the control group. Many studies in our review compared a specific PA intervention versus regular PE lessons or more PA versus less PA, often leading to little contrast in duration or frequency of PA between groups. When designing future studies, the contrast in PA between intervention and control groups needs careful consideration, specifically in the recruitment of schools (eg, striving for maximal contrast in PA but minimal differences with respect to background variables).

b. Most studies report on the ‘increased opportunities’ they offered children for being more physically active (eg, increase in the number of PE lessons per week). Few studies assessed the actual PA levels (eg, by accelerometer) or programme compliance (eg, attendance). Therefore, the exact dose of PA that children received is unknown, or it is unclear whether they compensated the increased PA in school with less PA after school. We recommend that future studies monitor actual PA levels and consider compliance in their analyses in addition to the standard intention-to-treat analysis.

c. Adequate sample sizes are needed to be able to pick up relevant changes in cognitive and academic performance, but also to reduce the likelihood of baseline differences between different conditions.

d. More effort should be directed towards the application of valid and reliable measures of cognitive performance using standardised and practical assessment tools, thereby creating better opportunities to compare findings across PA interventions more directly.

e. Many studies only mention whether the intervention effect was significant or not without providing information on effect sizes, confidence intervals or an exact P-value. To be able to judge not only the significance but also the practical relevance of the intervention effects, we recommend to always report both effect sizes and confidence intervals to enhance the value of the results presented in a single study and enable inclusion in future meta-analytic reviews.
CONCLUSIONS

Based on 11 high-quality intervention studies, we found inconclusive evidence of a beneficial effect of PA interventions on cognitive and overall academic performance in children. However, based on academic outcomes related to mathematics, we found strong evidence for beneficial effects of PA.

There is a need for more high-quality intervention studies to establish a potential causal relationship between PA and cognitive and academic performance. Many aspects of this possible relationship remain unclear: research priorities based on expert opinion highlight the need for future research into the causality of the ‘PA-cognitive performance’ relationship, including experimental studies on PA characteristics (eg, what type, duration, and intensity is most effective), studies on relevant moderators (eg, age, gender, socioeconomic status) and the underlying mechanisms.

By prioritising the most important research questions and formulating recommendations, we hope to guide researchers in generating high-quality evidence to further elucidate the relationship between PA and cognitive/academic performance.

What is already known?

► Many aspects of the potential relationship between PA and cognitive/academic performance in children remain unclear.
► Most reviews summarised the existing evidence at study level, not taking into account the number of outcomes analysed, thereby leading to a potential overestimation of the suggested relationship between PA and cognitive/academic performance.

What are the new findings?

► There is inconclusive evidence for a beneficial effect of PA interventions on cognitive and overall academic performance in children. There is strong evidence for beneficial effects of PA on maths performance.
► Research priorities based on expert opinion highlight the need for high-quality research into the causality of the ‘PA-cognitive performance’ relationship, including experimental studies on PA characteristics, relevant moderators, and underlying mechanisms.
► Expert opinion-based recommendations for future research are formulated with regard to study design, adequate control groups and sample size, measurement of cognitive performance and compliance, and data analysis.

References

Review

75 Hagins M, Rundle A. Yoga improves academic performance in urban high school students compared to physical education: a randomized controlled trial. Mind, Brain, Educ 2016;10:105–16.
84 Sjöwall D, Hertz M, Klingberg T. No long-term effect of physical activity intervention on working memory or arithmetic in preadolescents. Front Psychol 2017;8:1342.
90 Pesce C. Shifting the focus from quantitative to qualitative exercise characteristics in exercise and cognition research. J Sport Exerc Psychol 2012;34:766–86.