Physical activity and brain structure, brain function, and cognition in children and youth: A systematic review of randomized controlled trials

Katie E. Gunnell\textsuperscript{a,\textdagger}, Veronica J. Poitras\textsuperscript{b}, Allana LeBlanc\textsuperscript{c}, Kylie Schibili\textsuperscript{d}, Kheana Barbeau\textsuperscript{e}, Nina Hedayati\textsuperscript{f}, Matthew B. Ponitfex\textsuperscript{g}, Gary S. Gold\textsuperscript{f}\textdagger, Charlotte Dunlap\textsuperscript{b}, Emily Lehan\textsuperscript{i}, Mark S. Tremblay\textsuperscript{d}

\textsuperscript{a} Carleton University, Department of Psychology, Canada
\textsuperscript{b} Independent Researcher, Canada
\textsuperscript{c} Division of Prevention and Rehabilitation, University of Ottawa Heart Institute, Canada
\textsuperscript{d} Healthy Active Living and Obesity Research Group, Children’s Hospital of Eastern Ontario Research Institute, Canada
\textsuperscript{e} University of Ottawa, School of Psychology, Canada
\textsuperscript{f} Wilfred Laurier University, Department of Psychology, Canada
\textsuperscript{g} Michigan State University, College of Education, USA
\textsuperscript{i} University of Toronto, Canada

ARTICLE INFO

Keywords:
Brain health
BDNF
Executive function
Exercise
Achievement
Physical education

ABSTRACT

Background: Previous reviews show a favourable relationship between physical activity (PA) and brain health in children and youth. The purpose of this systematic review was to extend the generalizability of previous findings using only studies that employed randomized controlled designs in a wider age range.

Methods: After registration in PROSPERO, PRISMA guidelines were followed. Studies must have used a randomized controlled design; manipulated PA once (i.e., acute) or more (i.e., chronic) in apparently healthy children (1 month-17.99 years); and examined cognitive function, brain function, or brain structure as outcomes. Articles were reviewed for inclusion and data extraction were performed in duplicate.

Results: Overall, 84 studies from 83 papers with 12,600 unique participants were included (n\textsubscript{range} = 10–1,224, M\textsubscript{range} = 0.77–17 years). Studies were mostly rated as low risk of bias. The majority of studies (n = 52) reported at least one favourable cognitive function outcome associated with a PA intervention. Few studies (n = 6) reported an unfavourable outcome. Examining the multiple cognitive function outcomes within each study, PA had mostly no effect (n\textsubscript{acute} = 29, n\textsubscript{chronic} = 47) or a mix of favourable and no effects (n\textsubscript{acute} = 20, n\textsubscript{chronic} = 27).

For brain function, acute PA was associated with no change (n = 2) whereas chronic PA was associated with a mix of increases, decreases, or no change (n = 3). For brain structure, two overlapping studies found either favourable or no effects of chronic PA.

Discussion: PA is unlikely to harm brain health in children and youth and may confer some benefits. More research is needed to examine the relationship between PA and brain structure and function.

1. Introduction

Available evidence from recent systematic reviews and meta-analyses indicates that acute and chronic physical activity (PA) are positively correlated with brain health in children and youth (Álvarez-Bueno, Pesce, Cavero-Redondo, Sánchez-López, Garrido-Miguel, et al., 2017; Chang, Labban, Gapin, & Etnier, 2012; Ludgya, Gerber, Brand, Holsboer-Trachsler, & Pühse, 2016; Sibley & Etnier, 2003; Tomporowski, 2003a; Vazou, Pesce, Lakes, & Smiley-Oyen, 2016). In a recent position statement and systematic review of correlational and experimental research, Donnelly et al. (2016) concluded that PA in children aged 5–13 years is beneficial for cognition, brain function, and brain structure. Specifically, robust benefits were found when cognitive performance was measured by speed and accuracy assessments. Their results extended to indicators of academic achievement; however, results from experimental studies produced more equivocal findings (Donnelly et al., 2016). To date, the majority of previous research has focused on cognition, yet preliminary results from Donnelly and

\textsuperscript{\textdagger}Corresponding author. Carleton University, A511 Loeb Building, 1125 Colonel By Drive, Ottawa, ON, K1S 5B6, Canada.

E-mail address: Katie.gunnell@carleton.ca (K.E. Gunnell).

https://doi.org/10.1016/j.mhpa.2018.11.002
Received 6 April 2018; Received in revised form 10 November 2018; Accepted 12 November 2018
1755-2966/ © 2018 Elsevier Ltd. All rights reserved.
colleagues’ review show that physical fitness — and to a lesser extent overall PA — favourably modifies brain structures and functions that support cognitive control (also called executive functions). Indeed, researchers have hypothesized, and preliminary data suggest, that PA is a mechanism for favourable structural and functional changes to the brain (e.g., increased neurotrophins, cerebral blood flow or grey matter volume), which in turn, enhances cognitive outcomes (Lubans et al., 2016; Voss, Vivar, Kramer, & van Praag, 2013).

Previous systematic reviews on PA and brain health included correlational and experimental designs, largely focused on only cognitive function, and combined acute and chronic physical activity together. To extend the generalizability of previous systematic reviews, the purpose of our study was to conduct a systematic review of randomized controlled trials that examined the relationships between chronic and acute PA and indicators of cognitive function, brain function, and brain structure in a wide sample of children and youth. Cognitive function refers to the set of mental processes responsible for learning and understanding (Donnelly et al., 2016). Brain function refers to the functional neurological changes (e.g., expression of growth factors, neuronal activity) whereas brain structure refers to the structural neurological changes (e.g., white/grey matter volume) (Donnelly et al., 2016). Limitations of previous systematic reviews are apparent and provided the impetus for the current systematic review. First, many of the systematic reviews conducted to date have focused on a narrow age range (Donnelly et al., 2016; Venetsanou, Kambas, & Giannakidou, 2015). Although restricting age range may make a review more manageable, it reduces the generalizability of findings, may inhibit opportunities for knowledge translation, and may limit its use for public health purposes. Second, notwithstanding a few systematic reviews (Álvarez-Bueno, Pesce, Cavero-Redondo, Sánchez-López, Martínez-Hortelano, et al., 2017; Donnelly et al., 2016) many systematic reviews have included only one type of PA (e.g., hatha yoga; Luu & Hall, 2016; Lees & Hopkins, 2013) or only one main category of cognition (e.g., solely academic achievement; Álvarez-Bueno, Pesce, Cavero-Redondo, Sánchez-López, Garrido-Miguel, et al., 2017; or solely executive function; Luu & Hall, 2016). While these studies provide precise information for specific exposures and outcomes, a broad understanding of all PA exposures and cognitive outcomes has the largest potential to inform public health recommendations. Third, some systematic reviews have focused on specific sub-populations (e.g., only children with overweight or obesity; Bustamante, Williams, & Davis, 2016) limiting the inferences to all apparently healthy populations. Fourth, to date, most systematic reviews have relied on observational data (i.e., correlational studies), which limit their ability to explore causal relationships. Lastly, previous systematic reviews have included studies that may have been confounded by co-interventions. For example, if an intervention includes both PA and diet components and the control group did not also receive the diet component, it is difficult to know if the effects of the intervention are related to the PA exposure, the snack, or an interaction of the two.

Based on the limitations of previous systematic reviews outlined above, we sought to extend the literature while enhancing the generalizability of our findings for population health promotion. First, our review included all age ranges spanning 1 month to 17.99 years. This approach is similar to another recent systematic review (Vazou et al., 2016), yet extends their results to include both acute, and chronic PA interventions regardless of type, intensity, and/or duration. Second, we examined all brain-related outcomes spanning cognition, brain structure, and brain function. Third, we included all studies that examined an apparently healthy sample including, but not limited to, children with overweight and obesity. Fourth, we included only randomized controlled trials [RCT]. Finally, we ensured that included studies only manipulated PA or if another lifestyle behaviour was manipulated (e.g., snacks), it was manipulated identically across the control and intervention groups such that we could ensure the results were attributable strictly to PA.

## 2. Methods

We registered the systematic review through the International Prospective Register of Systematic Review (PROSPERO; CRD42016042116). The systematic review was conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Moher, Liberati, Tetzlaff, & Altman, 2009).

### 2.1. Eligibility criteria

The population, interventions, comparisons, outcomes, and study design framework (PICO: Schardt, Adams, Owens, Keitz, & Fontelo, 2007) was used to ensure study selection aligned with our research questions. Although not specified in our PROSPERO registration, studies were excluded if they examined active video gaming as their PA intervention. The decision to exclude these studies was made before results were summarized and because exposure to active video games could confound the association. Previous research with adults has shown an enhanced association between exergames and cognition compared to regular PA and cognition (Stanmore, Stubbs, Vancampfort, de Bruin, & Firth, 2017). Furthermore, from a public health perspective, exergames are not recommended as a method to increase overall PA or physical fitness (Chaput et al., 2013). Studies were also excluded if their results were presented by subgroups (e.g., if they categorized people into high fitness or low fitness) as it would have made it difficult to know the effect of PA on all participants in the RCT. Only RCTs were included in our systematic review. Cross-over studies where children served as their own control after a washout period were included if children and youth were randomized to begin with either the PA exposure or control. Studies that were quasi-experimental or non-experimental were excluded. Studies were included only if they were published in English or French. Published and in press peer-reviewed articles were included. Grey literature, theses, and conference abstracts or proceedings were excluded.

#### 2.1.1. Population

Apparently healthy children (general populations, including those with overweight/obesity, with no clinical diagnosis) aged 1 month-17.99 years.

#### 2.1.2. Intervention (Exposure)

PA. For the purpose of this review, PA was defined as any bodily movement that increases energy expenditure above resting rates (Caspersen, Powell, & Christenson, 1985). “Prone position” or “tummy time” was considered as PA exposure in infants (< 1 year). Total energy expenditure (as measured via doubly labelled water or direct/indirect calorimetry) and physical fitness were not considered proxies for PA. All interventions must have targeted PA exclusively and not multiple health behaviours (e.g., both PA and diet) unless the co-intervention was also applied to the control group.

#### 2.1.3. Comparator

Volume, duration, frequency, intensity, pattern (e.g., sporadic bouts, meeting/not meeting current guidelines), or type of PA.

#### 2.1.4. Outcome

Cognitive function, brain function, and brain structure. Intelligence and achievement tests were only included if they came from measures that are standardized (i.e., teacher evaluated grade point average, scores on non-standardized math tests etc. would have been excluded).

### 2.2. Search strategy

We developed an electronic search strategy in collaboration with a research librarian with expertise in systematic reviews. A second
research librarian peer-reviewed the search strategy. Using Ovid, we searched MEDLINE (1946-present), EMBASE (1980-to 2015 week 25; 1980–2017 week 31), Cochrane Central Register of Control Trials (February 2016 and June 2017), and PsycINFO (1806 to April week 1 2016; 1806 to July week 4 2017). EBSCOhost was used to search SPORTDiscus. The original searches were conducted in June 2016 and were updated in July 2017. Complete search strategies for each database and updated lines are provided in Supplementary File 1. Finally, a member of the research team checked reference lists from systematic reviews or protocols published in 2014 or later to identify other papers not captured in our search strategy.

All records were saved as text files from OVID and EBSCOhost and then exported to Reference Manager (Version 13; Thompson Reuters, San Francisco, CA). Duplicates were removed in Reference Manager. Titles and abstracts of all records were uploaded to DistillerSR (Evidence Partners; Ottawa, ON). After training, at least two independent reviewers reviewed all titles and abstracts against inclusion criteria. For a study to be excluded at the title and abstract level, both reviewers had to exclude it. For a study to pass to the next level of screening (i.e., full text), one reviewer was required to include it. At the level of screening full texts, two independent reviewers evaluated full text documents against inclusion criteria. Consensus on inclusion or exclusion was required before inclusion in the final study. Any conflicts between reviewers were resolved through discussion or by a third reviewer if needed.

2.3. Data extraction

After pilot testing the extraction forms, an author who had graduate statistics training at a minimum of a master’s degree level extracted the data from full texts into a table in Microsoft Word. A second independent reviewer with a PhD verified the accuracy of extraction. If changes were made, the same or a new reviewer with graduate training was asked to re-verify changes. The reviewers were not blind to the authors or journals during data extraction. The following information was extracted: Study characteristics including first author, publication year, journal, study design, country of study, sample size, age, and sex; PA intervention characteristics, including a description of treatments and control conditions (volume, frequency, duration, intensity, type); brain indicator (cognitive function, brain function, and brain structure); summary of findings and relevant descriptive statistics; and covariates used in the analysis, if applicable. Where applicable, results from the most adjusted models were extracted. We interpreted study findings as statistically significant at \( p < .05 \) even if the authors used other \( p \)-value thresholds. For the purposes of synthesis, a few rules of extraction were followed. In the case that an author reported a non-significant \( (p > .05) \) interaction but carried out other post-hoc analyses, only the interaction data were extracted and deemed null. In the case where authors did not include any information about testing interactions, within-group or between-group main effects were interpreted. Finally, in the case where authors did not report \( p \)-values, reported effect size thresholds were interpreted. Attempts were made to contact authors when discrepancies were found or if we were unsure how to interpret results as presented (see notes within Supplementary Files 2–6 where applicable).

2.4. Risk of bias and evidence quality assessment

Risk of bias was assessed using Cochrane’s risk of bias tool (Higgins & Green, 2011). We evaluated each study for risk of bias caused by inadequate random sequence generation; allocation concealment; blinding of participants and personnel, or blinding of outcome assessment; incomplete outcome data; selective reporting; and other sources of bias. The risk of bias was ranked as low, high or unclear and was supported by a description of why it was evaluated as such. Similar to the data extraction protocol outlined above, risk of bias assessment was conducted by one reviewer and verified by a second.

2.5. Data extraction and synthesis

Although many studies used the same measures (e.g., Flanker task), meta-analyses were not performed given the heterogeneity of the study designs, PA exposures (type, duration, frequency, etc.) and differential outcome reporting. Narrative synthesis was conducted and results were structured around (1) cognitive function, (2) brain function, and (3) brain structure. Results were further stratified by acute (i.e., only one PA bout) or chronic (i.e., repeated PA bouts over time) PA. Finally, specific results were further stratified based on the type of comparison being made: (1) PA compared to sedentary control, (2) multiple comparisons in one study with multiple exposures and/or controls, and (3) PA compared to another type, intensity, duration, volume, or pattern of PA. All cognition results were grouped according to the measure the authors used and in accordance with coding schemes used by other researchers (Chang et al., 2012; Pontifex et al., Submitted for publication). Examples of specific measures and their associated cognitive function category are available in Supplementary File 7. Outcome categories for cognitive function were:

1) Cognitive control (also called “executive function”) referred to the combination of processes employed when acting on instinct was insufficient or impossible (Diamond, 2013). Cognitive control included three sub-categories. Inhibition was the ability to control thoughts, attention, behaviour, and emotions while overriding urges and included tasks such as the Flanker task (Diamond, 2013). Working memory was the ability to hold information in the mind while working with it mentally and included tasks such as Digit Span (Diamond, 2013). Cognitive flexibility was the ability to shift approaches or perspectives to a challenge or change in rules and included tasks such as the Trail-making-test (TMT) (Diamond, 2013). Finally, unitary constructs was a sub-category that was used when a measure that assessed more than one sub-domain of cognitive control was used (e.g., the Tower of London task).

2) Attention represented the ability to maintain and resist distraction (Janssen, Toussaint, van Mechelen, & Verhagen, 2014b) and included tasks such as the d2 test and psychomotor vigilance task.

3) Information processing represented the identification of a stimulus, selecting an appropriate response, and responding (Tomporowski, 2003b) and included tasks such as visual search task and digit symbol substitution task.

4) Intelligence and achievement referred to intelligence tests (e.g., Eseck’s IQ Numerical Ability) and academic achievement tests related to education (e.g., reading, math).

5) Memory represented the ability to hold information in the mind (Diamond, 2013) and included tasks such as the delayed recall or paired association tasks. Memory is distinct from working memory (classified under cognitive control) because memory requires holding information in some capacity whereas working memory requires active maintenance and critically, the manipulation of information within memory (Cowan, 2008).

6) Motor speed and learning referred to the speed at which participants responded (e.g., choice response time) and the ability to acquire new knowledge (Donnelly et al., 2016).

7) Composite cognition was used when researchers used a combination of measures that spanned more than one of the categories above presenting the data in a manner such that separating the effects by aspects of cognition was not possible.

When there were multiple measures for the same outcome (e.g., Stroop and Flanker tasks used within one study), each outcome was summarized under only one overarching category (e.g., inhibition). For cognitive function and brain structure, data were synthesized for each category as “favourable” when all findings were favourable, “mixed
null and favourable” when a mix of null and favourable results were presented, “null” when only null results were presented, “unfavourable” when only unfavourable results were presented, “mixed null and unfavourable” when null and unfavourable results were presented, and “mixed favourable/unfavourable” when both favourable and unfavourable results were presented. For brain function a similar strategy was used; however, results were described as increased, decreased and/or no change. A secondary analysis was conducted to examine the results based on speed (e.g., assessments measured in milliseconds) and accuracy (e.g., assessments measured in number of errors, correct response) measures of cognitive function. For this analysis, only studies that separated speed or accuracy indicators were retained. Results were coded using the same outcome categories above, however each category was separated per assessment. For example, whereas inhibition was coded as only one outcome above, in this secondary analysis it was coded as inhibition speed and inhibition accuracy (if applicable; see Supplementary Files 8–9).

3. Results

3.1. Study descriptions

Results of the data screening are presented in the PRISMA flow diagram (see Fig. 1). Through the database searches, we identified 19,556 records; an additional seven studies were identified through reviewer nomination. After de-duplication, 17,220 records were retained for level 1 screening. After level 1 screening (titles and abstracts), 1078 records were retained to be screened in full. Of these, 83 records were identified as meeting inclusion criteria (see Fig. 1 for reasons of exclusion; a full list of excluded studies and reasons for exclusion is available in Supplementary File 10). Although 83 relevant records were identified, one study was split in to two because it presented two separate experiments (Soga, Shishido, & Nagatomi, 2015a; 2015b). As such, data below are presented for 84 studies.

In total, data for this review came from 12,752 participants across 84 studies (12,600 from 80 unique studies). Sample sizes ranged from 10 to 1224 and were extracted based on the smallest reported sample size (i.e., if the author reported sample size for pre-test and post-test accounting for missing data, the smaller sample size was used). Studies were included from 24 countries with the majority from the USA (n = 28) followed by Germany (n = 8), Australia, the United Kingdom (n = 6 each), India and Switzerland (n = 5 each), Italy (n = 4), Canada (n = 3), Netherlands, China, Denmark, Taiwan, Spain and Japan (n = 2 each), and Malaysia, Korea, South Africa, Philippines, Brazil, and Norway (n = 1 each). The reported mean age of the participants across all studies ranged from 0.77 to 17 years. Based on the mean age reported or the range/grade level if mean age was not reported, 58 (69%) studies included children (age 5–11.99 years), 20 (24%) studies included youth (age 12–17.99 years), 2 (2%) studies included preschoolers (age 1–4.99 years), 3 (4%) studies included both children and youth, and one (1%) study included infants (< 1 year). There were 83 studies that examined cognition (n = 12,734, nunique = 12,600 participants), 2 that examined brain structure (n = 36, nunique = 18 participants) and 8 that examined brain function (n = 207, nunique = 185 participants).

3.2. Cognitive Function

3.2.1. Cognitive control

3.2.1.1. Inhibition

3.2.1.1.1. Acute. Ten studies with 522 participants examined an acute bout of PA compared to a sedentary control (see Table 1). The majority of the studies (4/10) reported no effect (Cooper, Bandelow, Nute, Morris, & Nevill, 2012; Soga, Shishido, & Nagatomi, 2015b; Stein, Auerswald, & Ebersbach, 2017; Stroth et al., 2009) or mixed null and favourable effects (4/10) (Browne et al., 2016; Chen, Yan, Yin, Pan, &
Table 1
Narrative summary of the relationships between physical activity and cognitive control.

<table>
<thead>
<tr>
<th>PA Exposure</th>
<th>Acute</th>
<th>Chronic</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td># of participants (# of studies)</td>
<td>Absolute effect</td>
</tr>
<tr>
<td>PA vs None</td>
<td>522(10)</td>
<td>Inhibition</td>
</tr>
<tr>
<td></td>
<td>535(9)</td>
<td>Working Memory</td>
</tr>
<tr>
<td></td>
<td>613(5)</td>
<td>Cognitive Flexibility</td>
</tr>
<tr>
<td></td>
<td>392(3)</td>
<td>Multiple comparisons</td>
</tr>
<tr>
<td></td>
<td>276(2)</td>
<td>Working Memory</td>
</tr>
<tr>
<td></td>
<td>311(2)</td>
<td>Cognitive Flexibility</td>
</tr>
<tr>
<td></td>
<td>81(1)</td>
<td>Unitary Construct</td>
</tr>
</tbody>
</table>

4/10 studies reported mixed null and favourable effects (Brown et al., 2016; Chen et al., 2014; Cooper et al., 2016; Drollette et al., 2014)
4/10 studies reported null effects (Cooper et al., 2012; Soga et al., 2015a; Stein et al., 2017; Strath et al., 2009)
1/10 studies found mixed null and unfavourable effects (Soga et al., 2015a)
1/10 studies found favourable effect (Jäger et al., 2014)
2/9 studies found no effect (Cooper et al., 2016; Drollette et al., 2012; Howie et al., 2015; Jäger et al., 2014; Pietro, 1986)
1/9 studies found mixed null and unfavourable effect (Soga et al., 2015a)
2/5 studies found mixed null and favourable effects (Berse et al., 2015; Chen et al., 2014)
3/5 studies found null effects (Howie et al., 2015; Jäger et al., 2014; Stein et al., 2017)
1/3 studies found no effect (de Greeff et al., 2016; Krafft, Schwarz, et al., 2014; Torbeyns et al., 2017)
1/3 studies found mixed null and favourable effects (Cho et al., 2017)
1/3 studies found null or unfavourable effects (Tarp et al., 2016)
1/3 studies found no effect (de Greeff et al., 2016)
1/3 studies found a favourable effect (Leong et al., 2015)
1/3 studies found null and favourable effects (Hill et al., 2010)
1/3 studies found mixed null and unfavourable effects (Soga et al., 2015a)
1/2 studies found an unfavourable effect (Soga et al., 2015a)
1/2 studies found mixed null and favourable effects (Berse et al., 2015; Chen et al., 2014)
3/5 studies found null effects (Howie et al., 2015; Jäger et al., 2014; Stein et al., 2017)
1/2 studies found mixed null and favourable effects (S. Chen et al., 2016)
1/2 studies found no effect (de Greeff et al., 2016)
1/3 studies found mixed null and favourable effects (S. Chen et al., 2016)
1/3 studies found null or unfavourable effects (Jäger et al., 2014)
1/3 studies found null or unfavourable effects (Hill et al., 2010)
1/3 studies found mixed null and unfavourable effects (Soga et al., 2015a)
1/2 studies found no effect comparing low intensity exercise, high intensity exercise, or sedentary control (Budde et al., 2010)
1/2 studies found no effect between 4 conditions (physical games with cognitive engagement, aerobic games without cognitive engagement, cognitive games without physical activity, and sedentary control) (Jäger et al., 2015)
1/3 studies found (1) mixed-favourable effects for 30 min of physical activity compared to control (2) null effects of movement breaks compared to control (Kubesch et al., 2009)
1/2 studies found no effect comparing low intensity exercise, high intensity exercise, or sedentary control (Budde et al., 2010)
1/2 studies found no effect between 4 conditions (physical games with cognitive engagement, aerobic games without cognitive engagement, cognitive games without physical activity, and sedentary control) (Jäger et al., 2015)
1/2 studies found (1) mixed null and favourable effects between aerobic exercise and control and between aerobic and resistance, (2) null effects between resistance and control (Harveson et al., 2016)
1/2 studies found no effect between 4 conditions (physical games with cognitive engagement, aerobic games without cognitive engagement, cognitive games without physical activity, and sedentary control) (Jäger et al., 2015)
1/1 studies found no difference between 30 min of physical activity (continued on next page)
<table>
<thead>
<tr>
<th>PA Exposure</th>
<th>Acute</th>
<th>Chronic</th>
</tr>
</thead>
<tbody>
<tr>
<td># of participants (# of studies)</td>
<td>Absolute effect</td>
<td>Risk of bias</td>
</tr>
<tr>
<td>PA vs PA</td>
<td>26(1)</td>
<td>compared to control or between movement breaks compared to control (Kubesch et al., 2009)</td>
</tr>
<tr>
<td>Inhibition</td>
<td>1/1 studies found no effect of high intensity physical activity compared to low intensity physical activity (Chang et al., 2013)</td>
<td></td>
</tr>
<tr>
<td>PA vs None</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multiple comparisons</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PA vs. PA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>High</td>
<td>3/6 studies found favourable effects for: (1) developing motor skills compared to typical movement program (Robinson et al., 2016) (2) Taekwondo compared to regular physical education (Lakes et al., 2013), and (3) enriched physical education compared to regular physical education (Pesce et al., 2016).</td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>501(2)</td>
<td>2/6 studies found mixed null and favourable effects when comparing (a) enhanced physical education compared to typical physical education (Fisher et al., 2011) (b) mimicking actions of instructors when told to compared to regular activities involving physical exercise (Zhao et al., 2015)</td>
</tr>
<tr>
<td>Working Memory</td>
<td>487(3)</td>
<td>1/3 studies found no effect when comparing martial arts to standard physical education (Lakes &amp; Hoyt, 2004) or when comparing enriched physical education with regular physical education</td>
</tr>
<tr>
<td>Cognitive flexibility</td>
<td>501(2)</td>
<td>1/3 studies found favourable effects when comparing enhanced physical education compared to typical physical education (Fisher et al., 2011)</td>
</tr>
<tr>
<td>Low</td>
<td>472(3)</td>
<td>1/2 studies found mixed null and favourable results when comparing structured physical activity with unstructured physical activity such that structured performed better on some tasks but null effects were seen on others (Subramanian et al., 2015)</td>
</tr>
<tr>
<td>Unitary Construct</td>
<td>1/2 studies found a favourable effect for a running program compared to regular physical education (Tuckman &amp; Hinkle, 1996)</td>
<td></td>
</tr>
<tr>
<td>High</td>
<td>1/3 studies found a favourable effect for yoga compared to physical training (Manjunath &amp; Telles, 2001)</td>
<td></td>
</tr>
<tr>
<td>1/3 studies found mixed null and favourable effects for taekwondo compared to regular physical education (Lakes et al., 2013)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1/3 studies found no effect when comparing 325 min per week of physical activity compared to 135 min of physical activity per week (Kvalø et al., 2017)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Note:** PA = physical activity, “PA vs. None” represents studies that compared one physical activity treatment to some sedentary control condition. “Multiple comparisons” represents studies that had multiple interventions and/or control groups. “PA vs. PA” represents studies that compared a physical activity intervention to some other type/duration/intensity of physical activity.
Chang, 2014; Cooper et al., 2016; Drollette, Shishido, Pontifex, & Hillman, 2012). Of the studies that reported mixed null and favourable effects, no discernible pattern was noted. For example, one study reported null effects on reaction time and accuracy on congruent trials, but favourable effects on both during incongruent trials (Brownie et al., 2016). One study reported null effects on accuracy and favourable effects on reaction time (A. Chen et al., 2014). The remaining studies were mixed depending on the timing of measurement. One study reported no effect of accuracy during PA, but a favourable effect after the PA and no effect on reaction time during or after the PA (Drollette et al., 2012). Another study reported no effect on accuracy across outcomes but improvement in reaction time immediately following PA on the complex task and at follow-up of the simple task. However, no effects were seen in the assessments immediately following PA on the simple level or at follow-up on the complex level (Cooper et al., 2016). One study reported mixed null and unfavourable results whereby PA did not influence reaction time but was unfavourable for accuracy (Soga et al., 2015a). Finally, one study reported favourable results from PA (Jäger, Schmidt, Conzelmann, & Roebers, 2014).

Three studies with 392 participants (Harveson et al., 2016; Jäger, Schmidt, Conzelmann, & Roebers, 2015; Kubesch et al., 2009) had multiple comparisons and none found unfavourable effects (see Table 1). These studies found (1) favourable effects of aerobic exercise compared to seated control or resistance exercise, and no effect when comparing aerobic PA with resistance training (Harveson et al., 2016), (2) no differences between physical games with cognitive engagement, aerobic games without cognitive engagement, cognitive games without PA, and a sedentary control (Jäger et al., 2015b), and (3) mixed null and favourable effects of 30 min of PA compared to a control or no effects of movement breaks compared to the control (Kubesch et al., 2009). Finally, one study with 26 participants compared two types of PA and found no differences between high intensity PA and low intensity PA on inhibition (Chang, Tsai, Chen, & Hung, 2013) (see Table 1).

Of the studies that reported accuracy and speed separately, all compared PA to a sedentary control. Most reported null findings for accuracy (n = 6/9) (Chen et al., 2014; Cooper et al., 2012, 2016; Soga et al., 2015b; Stein et al., 2017; Stroth et al., 2009) followed by a mix of null and favourable (n = 2/9) (Brownie et al., 2016; Drollette et al., 2012) and unfavourable (n = 2/9) (Soga et al., 2015a) findings. Most reported null findings for speed (n = 6/9) (Cooper et al., 2012; Drollette et al., 2012; Soga et al., 2015b, 2015a; Stein et al., 2017; Stroth et al., 2009), followed by a mix of null and favourable (n = 2/9) (Brownie et al., 2016; Cooper et al., 2016) and favourable (n = 1/9) (Chen et al., 2014).

3.2.1.1.2. Chronic. Five studies with 1248 participants examined the effect of chronic PA compared to a sedentary control (see Table 1). The majority of the studies (3/5) reported null (de Greqff et al., 2016; Kraft, Schwartz, et al., 2014c; Torbeyns et al., 2017) effects. One study reported mixed null (Stroop word and Stroop colour) and favourable (Stroop colour-word) effects (Cho, So, & Roh, 2017). One study reported mixed null and unfavourable results whereby students who were assigned to continue regular schooling had more unfavourable reaction time interference scores when compared to students receiving 60 min of PA during the day and 5–10 min of physically active homework (Tarp et al., 2016). No differences were found on reaction time congruent or incongruent nor on accuracy congruent, incongruent, or interference (Tarp et al., 2016).

One study with 181 participants found no effect when comparing team games, aerobic exercise, and regular physical education (Schmidt, Jäger, Egger, Roebers, & Conzelmann, 2015b) (see Table 1). Six studies with 557 participants compared two conditions that both involved PA (see Table 1). Three of these studies found favourable effects (Lakes et al., 2013; Pesce et al., 2016; Robinson, Palmer, & Bub, 2016), with benefits seen in a motor skills program compared to typical movement program (Robinson et al., 2016), Taekwondo compared to regular physical education (Lakes et al., 2013), and enriched physical education compared to regular physical education (Pesce et al., 2016). Mixed null and favourable effects were shown in two studies. One compared enhanced physical education to typical physical education and found that reaction time did not change but accuracy improved (Fisher et al., 2011). Another compared an action mimicking task (i.e., Wesley/Simon says) with regular activity that involved PA and found that response inhibition was null for reaction time, omission errors, commission errors, inhibition hits and interference reaction time, incongruent accuracy, neutral accuracy, interference accuracy and favourable for incongruent reaction time, congruent reaction time, neutral reaction time, incongruent accuracy (Zhao, Chen, Fu, & Maes, 2015). No differences were found when comparing yoga with physical exercise (Telles, Singh, Bhardwaj, Kumar, & Balkrishna, 2013).

Of the studies that reported accuracy and speed separately, three compared PA to a sedentary control and two compared two types of PA. Of those comparing to a sedentary control, all reported null results for accuracy (n = 3/3) (Kraft, Schwarz, et al., 2014c; Tarp et al., 2016; Torbeyns et al., 2017) and null (n = 2/3) (Kraft, Schwartz, et al., 2014c; Torbeyns et al., 2017) or null and unfavourable (n = 1/3) (Tarp et al., 2016) for speed. Of those comparing two types of PA, accuracy was favourable (n = 1/2) (Fisher et al., 2011) or a mix of null and favourable (n = 1/2) (Zhao et al., 2015) whereas speed was null (n = 1/2) (Fisher et al., 2011) or a mix of null and favourable (n = 1/2) (Zhao et al., 2015).

3.3. Working memory

3.3.1. Acute

Nine studies with 535 participants examined an acute bout of PA and working memory (see Table 1). More than half (5/9) of the studies (Cooper et al., 2016; Drollette et al., 2012; Howie, Scharf, & Pete, 2015; Jäger et al., 2014; Pietro, 1986) found null results. Two studies found mixed null and favourable effects such that accuracy did not change but reaction time improved (A. Chen et al., 2014; Cooper et al., 2012). Another reported mixed null and unfavourable effects such that reaction time did not change but accuracy decreased (Soga et al., 2015a). Finally, in the same study, but different experiment, one experiment found unfavourable effects on reaction time and accuracy following PA (Soga et al., 2015b).

Two studies with 276 participants had multiple comparisons and both found null results when comparing low intensity, high intensity, and sedentary control groups (Budde et al., 2010) or four conditions with various combinations of PA and cognitive engagement (Jäger et al., 2015) (see Table 1).

Of the studies that reported accuracy and speed separately, all compared PA to a sedentary control. Most reported null results for accuracy (n = 4/5) (Chen et al., 2014; Cooper et al., 2012; Drollette et al., 2012; Soga et al., 2015a) followed by unfavourable (n = 1/5) (Soga et al., 2015b). For speed, favourable (n = 2/5) (Chen et al., 2014; Cooper et al., 2012) and unfavourable (n = 2/5) (Soga et al., 2015a; 2015b) were most reported followed by null (n = 1/5) (Drollette et al., 2012).

3.3.2. Chronic

Three studies with 1804 participants examined the effect of PA compared to a sedentary control (see Table 1). One study found no effect (de Greqff et al., 2016) and another found a favourable effect (Leong, Moghadam, & Hashim, 2015). The third study found mixed null and favourable effects where no effect was found on a size ordering task but favourable effects were found on listening and digit span for participants who performed PA in a second week during a cross-over study (Hill et al., 2010).

Only one study with 181 participants compared multiple groups and found no difference between team games, aerobic exercise, and regular physical education (Schmidt, Jäger, et al., 2015b) (see Table 1). Three
studies with 487 participants compared two types of PA (see Table 1). Two studies found no effect when comparing martial arts to standard physical education (Lakes & Hoyt, 2004) or when comparing enriched physical education with regular physical education (Pece et al., 2016). Another found favourable effects for enhanced physical education compared to typical physical education (Fisher et al., 2011).

3.4. Cognitive flexibility

3.4.1. Acute

Five studies with 613 participants examined an acute bout of PA and cognitive flexibility (see Table 1). Most of these studies (3/5) found null results with PA (Howie et al., 2015; Jäger et al., 2014; Stein et al., 2017). Two studies reported mixed null and favourable effects such that accuracy did not change but reaction time improved (A. Chen et al., 2014) or accuracy did not change but speed costs were more favourable with no change on speed and no switch trials (Berse et al., 2015).

Two studies with 311 participants had multiple comparisons (see Table 1). One of these studies found null results when comparing four conditions with various combinations of PA and cognitive engagement (Jäger et al., 2015). One study found mixed null and favourable results such that aerobic exercise compared to control had favourable improvements on the trail making test part B and TMTB-TMTA but no effect on TMTA (147). When comparing resistance exercise to the control, null results were seen on cognitive flexibility (Harveson et al., 2016). When comparing aerobic and resistance exercise, the aerobic group had more favourable scores on TMTB but no differences were seen in TMTA and TMTB-TMTA (Harveson et al., 2016).

Of the studies that reported accuracy and speed separately, all compared PA to a sedentary control. All reported null results for accuracy (n = 3/3) (Berse et al., 2015; A.; Chen et al., 2014; Stein et al., 2017) whereas there was null (n = 1/3) (Stein et al., 2017), favourable (n = 1/3) (Chen et al., 2014) and mix of null and favourable (n = 1/3) (Berse et al., 2015) for speed.

3.4.2. Chronic

No studies compared chronic PA to a sedentary control on cognitive flexibility (see Table 1). Two studies with 246 participants compared multiple conditions. One found null results when comparing aerobic exercise, resistance exercise, and regular physical education (Costigan, Eather, Plötz, Hillman, & Lubans, 2016). One study found first effects when comparing team games to regular physical education and team games to aerobic PA, and (2) null results when comparing aerobic PA to regular physical education (Schmidt, Jäger, et al., 2015b).

Two studies with 501 participants compared two PA groups (see Table 1). One study found mixed null and favourable results such that a structured PA group had more favourable scores on TMTB, Ruff Figural Fluency Test designs, rotations, and perseverations but no effect on TMTA compared to an unstructured PA group (Subramanian, Sharma, Arunachalam, Radhakrishnan, & Ramamurthy, 2015). Finally, one study found favourable effects such that participants in a running program performed better than those in a regular physical education program (Tuckman & Hinkle, 1986).

3.5. Unitary construct

3.5.1. Acute

One study with 81 participants examined the effect of acute PA compared to multiple comparisons. This study found no difference between 30 min of physical activity controlled to control or between movement breaks compared to control (Kubesch et al., 2009).

3.5.2. Chronic

Two studies with 549 participants examined the effect of chronic PA compared to sedentary control (see Table 1). One study found null results (de Greeff et al., 2016) and one found mixed null and favourable results such that total errors were favourable and nonperseverative errors were null on the Wisconsin Card Sorting Task (S. Chen, Tseng, Kuo, & Chang, 2016).

Three studies with 472 participants compared two types of PA (see Table 1). One study found favourable results for yoga compared to physical training in the Tower of London test (Manjunath & Telles, 2001). One study found mixed null and favourable results such that the Taekwondo group had more favourable scores on the Hearts and Flowers tasks of congruent accuracy, but no difference on congruent response time, incongruent accuracy, incongruent response time, mixed accuracy, and mixed response time (Lakes et al., 2013). The last study found null results when comparing 325 min of PA to 135 min of PA (Kvale, Bru, Brønnick, & Dyrdstad, 2017).

One study comparing two types of PA presented results separately for accuracy and speed. They found null results for speed and a mix of null and favourable for accuracy (Lakes et al., 2013).

3.6. Attention

3.6.1. Acute

Six studies with 571 participants examined the effect of an acute bout of PA compared to a sedentary control on attention (see Table 2). Three of the six studies found null results (Caterino & Polak, 1999; Ma, Le Mare, & Gurd, 2014; Mierau et al., 2014) and another two studies found mixed null and favourable results (Schmidt, Egger, & Gonzenbaum, 2015a; Tine & Butler, 2012). Of these mixed null results, one found no post-intervention effect for a cognitively demanding physical education lesson compared to regular school lessons (e.g., language lesson) but mixed favourable or null results 150 min later (Schmidt, Egger, et al., 2015a). Using the d2 test, the other found mixed favourable (total number of items correctly processed) and null (errors) results for PA compared to watching a movie (Tine & Butler, 2012). One study found a favourable effect of moderate to vigorous intensity PA compared to watching a movie (Niemann et al., 2013).

Four studies with 651 participants examined multiple comparisons (see Table 2). None of the results were unfavourable for the PA groups. One study found that attention improved over time in both low intensity and high intensity exercise groups, whereas both sedentary groups had decreased attention over time (Grieco, Jowers, Errisuriz, & Bartholomew, 2016). Another found improved attention scores after moderate and vigorous intensity movement breaks compared to no breaks and improved attention scores in the moderate intensity movement breaks compared to the vigorous intensity breaks (Janssen, Chinapaw, et al., 2014a). One found mixed null and favourable effects for (1) a cognitive treatment compared to physical exertion or sedentary control and (2) a PA plus cognitive engagement group compared to physical exertion of sedentary control (Schmidt, Benzing, & Kamer, 2016); however, no effect was found for PA compared to PA plus cognitive engagement, cognitive treatment, or sedentary control (Schmidt et al., 2016). Finally, one study did not provide enough information to determine where the simple effects were significant following the significant group by time interaction (Gallotta, Emerenziani, Franciosi, et al., 2015).

Only one study with 99 participants compared two PA conditions and found favourable effects such that attention scores improved after both coordinative exercise and regular sport lessons but the improvements were larger in the coordinative exercise condition (Budde, Voelcker-Rehage, Pietrašys-Kendziorra, Ribeiro, & Titow, 2008)(see Table 2).

Of the studies that reported accuracy and speed separately, three compared PA to a sedentary control, one had multiple comparisons, and one compared two types of PA. Of those that compared PA to a sedentary control, most reported null findings for both speed and accuracy (n = 2/3, n = 2/3, respectively; Ma et al., 2014; Mierau et al., 2014) or a mix of favourable and null (n = 1/3, n = 1/3, respectively;
<table>
<thead>
<tr>
<th>PA Exposure</th>
<th>Acute</th>
<th>Chronic</th>
</tr>
</thead>
<tbody>
<tr>
<td>PA vs None</td>
<td></td>
<td></td>
</tr>
<tr>
<td># of participants (# of studies)</td>
<td>Absolute effect</td>
<td>Risk of Bias</td>
</tr>
<tr>
<td>PA vs None</td>
<td>571(6)</td>
<td>3/6 studies found no effect (Caterino &amp; Polak, 1999; Ma et al., 2014; Mierau et al., 2014)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2/6 studies found mixed null and favourable effects (Schmidt, Egger, et al., 2015a; Tine &amp; Butler, 2012)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1/6 studies found a favourable effect (Niemann et al., 2013)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multiple comparisons</td>
<td>651(4)</td>
<td>1/4 studies found that both low intensity and high intensity exercise improved attention over time whereas two sedentary groups had decreases in attention over time (Gréco et al., 2016)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1/4 studies found improved attention scores after moderate and vigorous intensity movement breaks compared to no breaks and also improved attention scores in the moderate intensity movement breaks compared to the vigorous intensity breaks (Jansen, Chinapaw, et al., 2014a)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1/4 studies found mixed null and favourable effect for (1) a cognitive treatment compared to physical exertion or sedentary control and (2) a physical activity plus cognitive engagement group compared to physical exertion or sedentary control. No effect was found for physical activity compared to physical activity plus cognitive engagement, cognitive treatment, or sedentary control (Schmidt et al., 2016)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1/4 studies provided insufficient information to summarize the effect (Gallotta, Emerenziani, Franciosi, et al., 2015).</td>
</tr>
<tr>
<td>PA vs PA</td>
<td>99(1)</td>
<td>1/1 studies found improvements in attention following coordinative exercise and regular sport lessons but the coordinative exercise group had greater improvements (Bulde et al., 2008)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. PA = physical activity, "PA vs. None" represents studies that compared one physical activity treatment to some sedentary control condition. "Multiple comparisons" represents studies that had multiple interventions and/or control groups. "PA vs. PA" represents studies that compared a physical activity intervention to some other type/duration/intensity of physical activity.
Schmidt, Egger, et al., 2015a)). The study with multiple comparisons (Schmidt et al., 2016) found three null results for speed and two mixed null and favourable and one favourable for accuracy. The study that compared two types of PA (Budde et al., 2008) found one favourable effect for each speed and accuracy.

3.6.2. Chronic

Eight studies with 1809 participants examined the effect of chronic PA on attention compared to a sedentary control (see Table 2). Results from a pilot study indicated that movement-based learning integrated into math lessons had a mixed null and favourable effect, whereby attention scores improved at 6 weeks (post-test) but not at the midpoint of the intervention (3 weeks) (Riley, Lubans, Morgan, & Young, 2015). In the full intervention of this pilot study, the movement-based learning integrated into math lessons had a favourable effect on attention scores compared to regular math lessons (Riley, Lubans, Holmes, & Morgan, 2016). The majority of the studies (5/8) found no effect of PA compared to a sedentary control (Hill et al., 2010; Leong et al., 2015; Torbeyns et al., 2017; Webster, Wadsworth, & Robinson, 2015; Wilson, Olds, Lushington, Petkov, & Dollman, 2016). Finally, one study found favourable and unfavourable effects of PA such that under high demand, sport stacking had a favourable effect but under low demand, sport stacking had an unfavourable effect compared to an arts and crafts group (Mortimer, Krysztofik, Custard, & McKune, 2011).

Only one study with 156 participants had multiple comparisons and it found mixed favourable and null effects. In this study, a traditional PA group had faster processing speed compared to a control group but there were no differences on concentration and percent errors. The co-ordinative group had greater improvement in concentration and percent errors compared to the traditional and control groups but there were no difference in processing speed (Gallotta, Emerenziani, Iazzoni, et al., 2015) (see Table 2).

Finally, five studies with 757 participants compared two PA conditions (see Table 2). Two of five studies found favourable effects; martial arts had favourable effects compared to standard physical education (Lakes & Hoyt, 2004), and physical activities that emphasized directionality of movement (e.g., left-right, up-down, before-behind) had favourable effects compared to conventional physical education (Lipton, 1970). One study found mixed null and favourable effects such that scores on a two-target letter cancellation test were improved following structured PA, but commission or omission were unchanged (Subramanian et al., 2015). The remaining two studies found no differences between yoga and physical exercise (Telles et al., 2013) or taekwondo and regular physical education (Lakes et al., 2013).

Of the studies that reported accuracy and speed separately, two compared PA to a sedentary control and one compared two types of PA. In the two studies that compared PA to a sedentary control (Leong et al., 2015; Torbeyns et al., 2017), all results were null for both accuracy and speed. In the study that compared two types of PA (Subramanian et al., 2015), a null result was found for accuracy and a favourable result for speed.

3.7. Information processing

3.7.1. Acute

Two studies with 89 participants examined the effect of an acute bout of PA compared to a sedentary control on information processing (see Table 3). One study found no effect on information processing when comparing high intensity intermittent running to seated rest (Cooper et al., 2016). The other study found mixed favourable and unfavourable results when comparing a multi-stage fitness test to resting, such that reaction time scores were favourable but accuracy scores were unfavourable (Cooper et al., 2012).

<table>
<thead>
<tr>
<th>PA Exposure</th>
<th>Acute</th>
<th>Chronic</th>
<th>Risk of bias</th>
</tr>
</thead>
<tbody>
<tr>
<td>PA vs None</td>
<td>89(2)</td>
<td>8/4(3)</td>
<td>Low</td>
</tr>
<tr>
<td>Multiple comparisons</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>PA vs PA</td>
<td>488(2)</td>
<td>3/3(1)</td>
<td>High</td>
</tr>
</tbody>
</table>

Note: PA = physical activity, PA vs. None = compare one physical activity treatment to some sedentary control condition. “Multiple comparisons” represents studies that had multiple interventions and/or control groups. “PA vs. PA” represents studies that compared a physical activity intervention to some other type/duration/intensity of physical activity. Studies with the same number of * contain overlapping participants.
3.7.2. Chronic

Five studies with 1659 participants (four unique studies with 1641 unique participants) examined the effect of chronic PA compared to a sedentary condition on information processing (see Table 3). One study found a favourable effect when comparing PA to a control condition that was not described (Reed et al., 2010), whereas four studies (three unique) found no effects (Hill et al., 2010; Krafft, Pierce, et al., 2014a; Krafft, Schaeffer, et al., 2014c; Riley et al., 2016).

Two studies with 265 participants (one unique study with 77 unique participants) had multiple comparisons and found mixed favourable and null results (see Table 3). In one study, participants in the high dose exercise group performed better on Cognitive Assessment Scale planning compared to low dose exercise and control groups, whereas null effects were seen for other Cognitive Assessment Scale subscales. In this same study, the low dose exercise group did better on Cognitive Assessment Scale planning compared to the control group but showed no difference on all other Cognitive Assessment Scale subscales were null (Davis et al., 2011). Another study with the same sample found (1) no differences between high dose and low dose PA groups on Cognitive Assessment Scale subscales, (2) favourable effects on planning when comparing high dose to control but null effects on other Cognitive Assessment Scale subscales, (3) no effects when comparing low dose to control, and (4) null effects when comparing the combined low dose and high dose group with the control group (Davis et al., 2007).

Three studies with 448 participants compared two PA conditions and found null results between running and regular physical education (Tuckman & Hinkle, 1986), enriched physical education with regular physical education (Pesce et al., 2016), or enhanced physical education with standard physical education (Fisher et al., 2011) (see Table 3).

Of the studies that reported accuracy and speed separately, one compared PA to a sedentary control. In this study (Cooper et al., 2012), accuracy was unfavourable and speed was favourable.

3.8. Intelligence and achievement tests

3.8.1. Acute

Only one study with 17 participants compared multiple conditions involving an acute bout of PA (see Table 4). This study found moderate intensity exercise had null (for sentence comprehension), favourable (for spelling and reading), and unfavourable (for math) results on intelligence and achievement tests compared to seated rest (Duncan & Johnson, 2014). Vigorous and moderate intensity exercise groups combined had null (for reading, sentence comprehension), favourable (spelling), and unfavourable (arithmetic) results on intelligence and achievement compared to seated rest. Moderate intensity exercise had null (for spelling, reading, arithmetic, and sentence comprehension) results on intelligence and achievement compared to vigorous exercise (Duncan & Johnson, 2014).

3.8.2. Chronic

Nine studies with 2202 participants compared the effect of chronic PA on intelligence and achievement tests compared to a sedentary control condition (see Table 4). The majority of these studies (5/9) found null results (Ahamed et al., 2007; Donnelly et al., 2017; Fedewa, Ahn, Erwin, & Davis, 2015; Tarp et al., 2016; Torbeyns et al., 2017). One study found that infants who were given passive cycling exercise alternating with rest had favourable language scores compared to infants without the exercise manipulation (Porter, 1972). Finally, three studies found mixed null and favourable results when comparing PA to control conditions (Reed et al., 2010; Santos, Jiménez,ampaio, & Leite, 2017; Uhrich & Swalm, 2007). In one study, physical activity integrated into core curriculum had favourable (social studies) and no effect (English, Arts, Mathematics, Science) compared to a control that was not explicitly defined but assumed to be the regular curriculum (Reed et al., 2010). In another study, children in a sport stacking intervention (i.e., an activity that requires cross-lateral movements with both hands) had improved reading comprehension and no change in reading decoding compared to a group that had snack time (Uhrich & Swalm, 2007). Finally, participants in an extracurricular training program had favourable scores on a composite creative thinking score, elaboration, originality, closure, and titles but no change on fluency compared to a control group that was not described (Santos et al., 2017); although caution is warranted as these results were based on a magnitude inference approach and not statistical significance.

Four studies with 1141 participants had multiple comparisons (see Table 4), most of which were mixed null and favourable (3/4) with one showing mixed null, favourable, and unfavourable (1/4). One study found mixed-favourable results when comparing increased intensity with control and increased intensity plus increased number of physical education classes with control. Null results were found when comparing increased number of physical education classes with control (Ardoy et al., 2014). Another study found mixed favourable results when comparing high dose vs. low dose PA and null results when comparing high dose vs control and low dose vs. control (Davis et al., 2011). Another study found mixed-favourable results when comparing gross motor math with fine motor math and null effects when comparing gross motor math to control and fine motor math to control (Beck et al., 2016). Finally, one study found mixed favourable effects of the intervention in one cohort and mixed favourable and unfavourable effects in a second cohort. In cohort 1, students in the specialist trained teacher condition improved reading compared to the control, whereas those in the teacher trained condition improved language compared to the control, but null effects were seen for math or the basic battery score. In cohort 2, students with the trained teacher had improved reading and basic battery scores compared to specialist or control. Students with a specialist had unfavourable language scores compared to teacher trained or control. No effects were seen for math (Sallis et al., 1999).

Five studies with 546 participants compared two PA conditions (see Table 4). Most studies (3/5) found favourable results such that martial arts improved more than standard physical education (Lakes & Hoyt, 2004), integrated physical education with music improved more than movement exploration (Brown, Sherrill, & Gench, 1981), and enhanced physical education that emphasized directionality improved more than traditional physical education (Lipton, 1970). One study found mixed-favourable results, whereby fluency was improved more following aerobic PA than following traditional physical education (Herman-Toffler & Tuckman, 1998). One study found null results when comparing yoga to stretching and aerobic exercise (Chaya, Nagendra, Selvam, Kurpad, & Srinivasan, 2012).

3.9. Memory

3.9.1. Acute

One study with 70 participants examined the impact of an acute bout of PA compared to a sedentary control and found null results (Pietro, 1986) (see Table 5). Two studies with 139 participants examined multiple comparisons and found mixed null and favourable results for circuit training vs. control, or team games vs. control (Pesce, Crova, Cereatti, Castella, & Bellucci, 2009), and null results for any group comparing integrated learning and PA, non-integrated learning and PA, or standing (Mavilidi, Okely, Chandler, & Paas, 2016). Finally, one study with 43 participants compared two PA conditions and found that running had mixed favourable and null results on memory compared to a light PA and stretching group (Etier, Labban, Piepmeier, Davis, & Henning, 2014).

Of the studies that reported accuracy and speed separately, one compared PA to a sedentary control and found null results for both speed and accuracy (Pietro, 1986).

3.9.2. Chronic

One study with 44 participants examined the impact of chronic PA compared to a sedentary control and found null results (Torbeyns et al.,..
Table 4
Narrative summary of the relationships between physical activity and intelligence and achievement tests.

<table>
<thead>
<tr>
<th>PA Exposure</th>
<th>Acute</th>
<th>Chronic</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td># of participants (# of studies)</td>
<td>Absolute effect</td>
</tr>
<tr>
<td>PA vs None</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multiple comparisons</td>
<td>17(1)</td>
<td>1/1 studies found moderate exercise had null, favourable, and unfavourable effects on intelligence and achievement tests compared to seated rest. Vigorous and moderate exercise groups combined had null, favourable, and unfavourable effects on intelligence and achievement compared to seated rest. Moderate exercise had null effects on intelligence and achievement compared to vigorous (Duncan &amp; Johnson, 2014)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PA vs PA</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. PA = physical activity, “PA vs. None” represents studies that compared one physical activity treatment to some sedentary control condition. “Multiple comparisons” represents studies that had multiple interventions and/or control groups. “PA vs. PA” represents studies that compared a physical activity intervention to some other type/duration/intensity of physical activity.
Table 5
Narrative summary of the relationships between physical activity and memory.

<table>
<thead>
<tr>
<th>PA Exposure</th>
<th>Acute</th>
<th>Chronic</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td># of participants (# of studies)</td>
<td>Absolute effect</td>
</tr>
<tr>
<td>PA vs None</td>
<td>70(1)</td>
<td>1/1 study found no effect (Pietro, 1986)</td>
</tr>
<tr>
<td>Multiple comparisons</td>
<td>139(2)</td>
<td>1/2 studies found mixed null and favourable effects for circuit training vs. control or team games vs. control (Pesce et al, 2009)</td>
</tr>
<tr>
<td></td>
<td>43(1)</td>
<td>1/1 study found that the intervention group (running) had mixed null and favourable effects on memory compared to a control group (light physical activity and stretching) (Etnier et al, 2014)</td>
</tr>
</tbody>
</table>

Note. PA = physical activity, PA vs. none represents studies that compared one physical activity treatment to some sedentary control condition. Multiple comparisons represents studies that had multiple interventions and/or control groups. PA vs. PA represents studies that compared a physical activity intervention to some other type/duration/intensity of physical activity.

Table 6
Narrative summary of the relationships between physical activity and motor speed and learning.

<table>
<thead>
<tr>
<th>PA Exposure</th>
<th>Acute</th>
<th>Chronic</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td># of participants (# of studies)</td>
<td>Absolute effect</td>
</tr>
<tr>
<td>PA vs None</td>
<td>72(1)</td>
<td>1/1 studies found mixed null and favourable effects (Ellemberg &amp; St. Louis-Deschênes, 2010)</td>
</tr>
<tr>
<td>Multiple comparisons</td>
<td>126(2)</td>
<td>1/1 studies found favourable effects such that those in a contralateral ball bouncing intervention improved contralateral reaction time more than those in a control or ipsilateral ball bounding intervention. Children in the ipsilateral ball bouncing intervention improved their ipsilateral reaction time more than those in the contralateral ball bounding intervention and control group (Pedersen, 2014).</td>
</tr>
<tr>
<td></td>
<td>–</td>
<td>1/1 studies found no difference between 30 min of physical activity compared to control or between movement breaks compared to control (Kubesch et al., 2009)</td>
</tr>
<tr>
<td></td>
<td>–</td>
<td>1/2 studies found null effects when comparing sport stacking with standard physical education (Hart et al, 2005)</td>
</tr>
</tbody>
</table>

Note. PA = physical activity, “PA vs. None” represents studies that compared one physical activity treatment to some sedentary control condition. “Multiple comparisons” represents studies that had multiple interventions and/or control groups. “PA vs. PA” represents studies that compared a physical activity intervention to some other type/duration/intensity of physical activity.
2017) (see Table 5).

3.10. Motor speed and learning

3.10.1. Acute

One study with 72 participants examined the impact of an acute bout of PA compared to a sedentary control and found mixed null and favourable results (Ellemberg & St-Louis-Deschênes, 2010) (see Table 6). In this study, PA had a favourable effect on reaction time and choice response time, but there was no effect on choice response time accuracy (Ellemberg & St-Louis-Deschênes, 2010). Two studies with 126 participants had multiple comparisons. One found favourable effects such that children who were in contralateral and ipsilateral ball bouncing conditions improved reaction time in their respective directions (Pedersen, 2014). The other found no difference between 30 min of physical activity compared to control or between movement breaks compared to control (Kubesch et al., 2009).

One study compared PA to a sedentary control analyzed accuracy and speed separately and found null results for accuracy and favourable results for speed (Ellemberg & St-Louis-Deschênes, 2010).

3.10.2. Chronic

Two studies with 508 participants compared the effect of two PA interventions on motor speed and learning (see Table 6). One study found a favourable effect, whereby motor speed and learning improved following both structured and unstructured physical activities with greater improvement following structured physical activities (Sharma et al., 2015). In another study, no effect was found when comparing sport stacking with standard physical education (Hart, Smith, & DeChant, 2005).

3.11. Cognition composite

3.11.1. Chronic

Three studies with 1794 participants examined the effect of PA compared to sedentary control on measures that represented a composite of cognition categories (see Table 7). One study found null results (Kraft, Schaffer, et al., 2014c); however, two studies found mixed null and favourable effects (Hill et al., 2010; Hill, Williams, Aucott, Thomson, & Mon-Williams, 2011). In both of these latter studies, children who engaged in PA in week 2 had more favourable scores compared to controls, whereas those who engaged in PA in the first week had no difference in scores (Hill et al., 2010, 2011).

3.12. Brain function

3.12.1. Acute

Two studies with 43 participants examined brain activation after an acute bout of PA (see Table 8). One study found no change in brain activation (Stroth et al., 2009) and the other found a mix of increased activation (Δalpha-1 power increased during eyes open at rest), no change (Δalpha-1 power, eyes closed at rest) and decreased activation (Δbeta-1 and Δbeta-2 at rest and during a cognitive task) (Mierau et al., 2014). One study with 25 participants compared moderate and low intensity PA and found no change between groups in brain activation over time (Chang et al., 2013).

3.12.2. Chronic

Three studies with 86 participants examined brain activation after chronic PA interventions (see Table 8). Most of these studies (2/3) found mixed increases, decreases and no change in brain activation. Of these, one found that activation increased during a cognitive control task (i.e., antisaccade) and decreased or showed no change during another cognitive control task (i.e., Flanker) (Kraft, Schaffer, et al., 2014c). Another found no change in frontal eye fields or supplementary eye fields, increases in bilateral prefrontal cortex and decreases in...
<table>
<thead>
<tr>
<th>PA Exposure</th>
<th>Acute</th>
<th>Chronic</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td># of participants (# of studies)</td>
<td>Absolute effect</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PA vs None</td>
<td>43(2) Brain activation 1/2 studies found no change in brain activation (Stroth et al., 2009)</td>
<td>1/2 studies found mixed findings such that Δalpha-1 power increased during eyes open and did not change with eyes closed. Δbeta − 1 decreased and Δpower-2 decreased at frontal sites during rest. When comparing eyes open at rest to during a cognitive task, Δbeta − 1 and Δbeta − 2 decreased more in the frontal sites and Δbeta − 2 decreased more in the central sites (Mierau et al., 2014)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Risk of bias</td>
</tr>
<tr>
<td>PA vs PA</td>
<td>26(1) Brain Activation 1/1 studies found no change between moderate intensity and low intensity PA (Chang et al., 2013)</td>
<td></td>
</tr>
</tbody>
</table>

**Note.** PA = physical activity, “PA vs. None” represents studies that compared one physical activity treatment to some sedentary control condition. “Multiple comparisons” represents studies that had multiple interventions and/or control groups. “PA vs. PA” represents studies that compared a physical activity intervention to some other type/duration/intensity of physical activity.
bilateral posterior parietal cortex (Davis et al., 2011). The other study found no change in activation (Torbeys et al., 2017). One study with 37 participants examined resting state synchrony and found a mix of increases, decreases, and no change (Krafft, Schaeffer, et al., 2014c). One study with 30 participants examined brain derived neurotrophic factor and found increases after PA (Cho et al., 2017). In that same study, the authors also examined cerebral blood flow velocity and found no changes (see Table 8).

3.13. Brain structure

Two studies with 18 participants from the same RCT examined white matter integrity (see Table 9). One study found favourable effects such that participants in the chronic PA condition had improved white matter structural coherence and myelination compared to the control group (Schaeffer et al., 2014). The other study with the same sample found no results (Krafft, Schaeffer, et al., 2014c).

3.14. High-level summary and risk of bias

Detailed risk of bias appraisals are presented in Tables 1–9 Collectively, for acute PA, the majority of outcome categories (65%) were rated as low risk of bias followed by high (20%) or unclear (15%). For chronic PA, the majority (66.7%) of outcome categories were rated as low risk of bias followed by high (33.3%). The most common source of bias rated as “high” was selective reporting such that authors frequently reported results as figures only (i.e., did not provide means, standard deviations etc.) or did not report data from the non-significant results. For cognitive function, results from acute PA bouts were mostly null (n = 33), a mix of null and favourable (n = 20) or favourable (n = 12) with seven findings representing a mix of results that included at least one unfavourable relationship (see Fig. 2). Interpreting each study as a whole rather than individual outcomes within each study, the majority (n = 21/34) reported at least one favourable outcome and few reported at least one unfavourable effect (n = 3/34). For chronic PA interventions, results were mostly null (n = 48), a mix of null or favourable (n = 27), or favourable (n = 18) with three findings containing at least one unfavourable outcome (see Figs. 2 and 3). Interpreting each study as a whole rather than individual outcomes within each study, the majority (n = 31/49) reported at least one favourable outcome and few reported at least one unfavourable effect (n = 3/49).

Visually examining all cognitive function categories together for the acute studies wherein the authors separated accuracy and speed, results demonstrated that there was a higher frequency of favourable results for speed compared to accuracy and more null results for accuracy compared to speed (see Fig. 4). Visually examining all cognitive function categories together for the chronic studies wherein the authors separated accuracy and speed, results demonstrated similar patterns across both outcomes (see Fig. 4).

For brain function, results from acute PA bouts were no change (n = 2), or a mix of increases, decreases, and no change (n = 1). Results from chronic PA interventions were associated with increases (n = 1), no change (n = 2), or a mix of increases, decreases, and no change (n = 3). For brain function, most authors interpreted their results as suggestive that increases, decreases, or no change as a result of PA supported brain function (see Figs. 2 and 5). For brain structure, results of chronic PA interventions were favourable (n = 1) and null (n = 1); however, both of these findings came from the same study (see Figs. 2 and 5).

4. Discussion

In this study, we examined the relationships between PA and cognitive function, brain function, and brain structure in children and youth aged 1 month to 17.99 years old. We included 84 studies from 83 papers representing 12,600 unique participants. Results indicated that the studies were mostly rated as having low risk of bias. Overall, given that very few studies demonstrated unfavourable outcomes and most demonstrated no effect, it appears that PA is unlikely to be harmful and in some cases may be beneficial for cognitive function in children and youth. Given the small number of studies in limited age groups on brain function and structure, more research is needed to determine the effect of PA.

With respect to cognitive function, our results demonstrated that over half of the studies reported at least one favourable outcome. Notwithstanding, when each outcome was considered within each study, the results indicated that most of the results were null or a mix of null and favourable. These findings are somewhat inconsistent with past systematic reviews that found relatively consistent favourable effects (Donnelly et al., 2017). There are several explanations for these differences. First, our study excluded many RCTs that came from the Fitness Improves Thinking in Kids (FITKids) trials (Hillman et al., 2014) and other high quality studies that included multiple behavioural components. The studies emanating from the FITKids trials did not meet our inclusion criteria because they involved a nutrition and educational component whereas no such components were delivered in the waitlist control group. Therefore, it would have been difficult to determine if any intervention effects were attributable to PA alone or the combination of PA, nutrition, and education. Second, it is possible our results are slightly less favourable because we examined the frequency of each outcome within each study separately rather than summarizing the results based on at least one favourable effect in each study. Third, our review included more RCTs that have not been located in previous systematic reviews indicating that our sources are different. Despite these discrepancies, when these results are considered alongside previous systematic reviews and meta-analyses, the evidence remains in favour of promoting PA for children’s brain health.

Our results extend findings from a recent position statement and systematic reviews that examined these relationships with narrower parameters in terms of age range, exposure, use of correlational studies,
or outcomes (Chang et al., 2012; Janssen, Chinapaw, et al., 2014a; Mura, Vellante, Egidio Nardi, Machado, & Giovanni Carta, 2015; Vazou et al., 2016; Verburgh, Königs, Scherder, & Oosterlaan, 2013). Generally, we found evidence that acute and chronic PA has either no effect or favourable effects on cognitive function, brain function, and brain structure. Cognitive control was the most frequently examined outcome following an acute bout of PA, whereas indicators of intelligence and achievement tests were the most frequently examined outcomes following chronic PA. In both acute and chronic interventions, memory, motor speed, and learning were least frequently studied. There is a clear need for more research examining memory, motor speed and learning. Additionally, although results from chronic PA RCTs were mostly null or mixed null and favourable for information processing, there were very few studies examining an acute bout of PA. Researchers should further explore the relationships between acute PA and information processing to determine if it has similar effects on this outcome as what was reported after exposure to chronic PA.

Our results were generally consistent across acute and chronic PA exposures and aligned with previous literature (Chang et al., 2012; Donnelly et al., 2016; Luu & Hall, 2016; Vazou et al., 2016; Verburgh et al., 2013) to suggest that PA is beneficial, or at least does not impede inhibition, cognitive flexibility, or working memory. Interestingly, a few previous meta-analyses have shown no significant effect of acute PA on working memory in children and adults or young adults (Chang et al., 2013; Verburgh et al., 2013). Our results suggest largely null findings across acute and chronic PA exposures with respect to working memory with the exception of a few unfavourable effects. Donnelly et al. (2016) have suggested that the effect of PA on cognitive control might be dependent on the type of PA exposure. To this end, Vazou et al. (2016) found evidence for enhanced improvements following PA with skills training or cognitive engagement compared to simple aerobic activities. We found no discernible pattern for our mixed results. Future research is warranted to further delineate which aspects of PA influence working memory in children and youth.

Researchers have shown that participants make tradeoffs when using speed and accuracy assessments (Wickelgren, 1977). It is possible that children and youth prioritized responding differently across these studies (Donnelly et al., 2016). Nonetheless, a recent meta-analysis examining the effect of acute PA on executive function (or cognitive control) in children and adults noted no such tradeoff effect (Ludyga, Gerber, Brand, Holsboer-Trachsler, & Pühse, 2016). In this study, when considered overall, the unfavourable effects were always, except in one study (Soga et al., 2015b), accompanied by favourable or null effects on other outcomes in that same domain. The secondary analysis examining the frequency of results based on speed or accuracy demonstrated that visually, there was a higher frequency of favourable findings for speed compared to accuracy after acute PA. Further, there were more null findings on accuracy compared to speed after acute PA. This finding is novel and should be further examined using larger sample sizes such that statistical analysis can be carried out rather than examining simple frequencies. No discernible visual pattern emerged when examining speed and accuracy after chronic PA. Researchers may wish to systematically elucidate if certain types of chronic PA are conducive to improving speed versus accuracy or vice versa or if chronic PA has a more general rather than specific effect on cognitive function (Ludyga et al., 2016).

The effect of an acute bout of PA on attention, memory, and motor speed and processing was consistent with expectations and
Fig. 3. High-level summary of the relationships between acute and chronic PA and cognitive function per outcome.
Note. F = favourable, NF = null-favourable, FU = favourable-unfavourable, FUN = favourable, unfavourable, null, U = unfavourable, insuf = insufficient detail. Frequencies of each finding do not add up to the number of studies given that some studies reported multiple results when comparing multiple arms of the intervention.

Fig. 4. High-level summary: The effect of physical activity on speed and accuracy assessments of cognitive function.
Note. F = favourable, NF = null-favourable, N = null, U = unfavourable, NU = null-favourable. Frequencies of each finding do not add up to the number of studies given that some studies reported multiple results when comparing multiple arms of the intervention.
demonstrated mostly a mix of favourable and null effects. Another systematic review in children 4–18 years old focusing on attention after an acute bout of PA found inconsistent and weak results (Janssen, Chinapaw, et al., 2014a). Our study extends their review with additional higher quality studies and points towards the beneficial effects of both acute and chronic PA for attention with only one unfavourable association noted. Only a few studies identified herein examined memory and more research is needed to further elucidate the impact of acute and chronic PA on memory in children and youth.

Given that we separated studies that used an active control group (e.g., traditional physical education or some other PA exposure) from those that used a sedentary control, we were able to show that no RCT wherein PA was compared against another type of PA exposure produced unfavourable effects. Results from a previous meta-analysis with chronic PA interventions demonstrated that the strongest effects were seen between PA and sedentary control and the pooled effects were weakened or null when compared against another type of PA (Vazou et al., 2016). Based on our results that included acute exposures and this previous systematic review, it seems reasonable to suggest that both chronic and acute PA are beneficial, or at least not detrimental to cognition when compared to other types of PA.

When acute PA was examined, results suggested that PA supported brain activation needs for cognitive processes. Comparatively more studies examined brain function following chronic PA. Many of these studies have been identified by previous reviews (Donnelly et al., 2016; Mura et al., 2015), yet our review included an additional study that found benefits in brain-derived neurotrophic factor following chronic PA and no change in cerebral blood flow velocity. Overall, it appears that acute and chronic PM may be beneficial for activating regions of the brain and biomarkers that are necessary for cognitive control; however, caution is warranted when interpreting these results given that there were few studies and they came from a narrow age range. Nonetheless, there is a need for more research examining brain function and structure in children and youth.

4.1. Practical implications

To the best of our knowledge, this is the first systematic review to examine the relationships between PA and cognitive function, brain function, and brain structure that is generalizable to all apparently healthy children and youth < 18 years old. Our results from 84 randomized controlled studies examining chronic and acute bouts of PA suggest that PA is associated with favourable or at least not detrimental effects on cognitive function, brain function and brain structure. Our results align with previous reviews that focused on more narrow age ranges, exposures, or outcomes (Bustamante et al., 2016; Chang et al., 2012; Donnelly et al., 2016; Norris, Shelton, Dunsmuir, Duke-Williams, & Stamatakis, 2015). In only a few instances, unfavourable effects were demonstrated. Nonetheless, these unfavourable effects were sparse and when considering the results as a whole, we do not believe they provide sufficient justification to not recommend the inclusion of PA for healthy brain development in children and youth. Results from our review support and extend previous conclusions (Álvarez-Bueno, Pesce, Cavero-Redondo, Sánchez-López, Garrido-Miguel, et al., 2017; Donnelly et al., 2016; Mura et al., 2015) that PA incorporated into school settings or elsewhere is unlikely to negatively impact brain health and outcomes related to academic achievement and intelligence test scores and may even be beneficial.

4.2. Limitations and future directions

There are numerous potential confounds in physical activity interventions that make it difficult to truly understand the independent effect of PA on brain health. Our attempt to isolate the effect of PA on brain health by excluding studies with multiple distinct behavioural components could be considered a limitation of this study since PA is not independent of other behaviours in the real world (e.g., an increase in PA may displace sedentary time, or may influence dietary habits). It is also possible that the studies that met inclusion criteria unintentionally influenced other behaviours, physiological factors, or psychosocial factors thereby confounding the relationships between PA and brain health. For instance, we included studies that integrated education and PA (e.g., children were active while learning). In other studies, interventions included musical components or cooperative skills during PA that could have confounded the effects of PA on brain health. Finally, we included studies that used a waitlist control where there was no attempt by the researchers to control for confounding factors in the control group. Therefore, despite excluding studies that explicitly involved multiple distinct behavioural components, it is possible that other factors confounded the relationships observed in this
study. Given the large heterogeneity of our included studies (see Tables 1–9) meta-analyses was ill-advised and results are based on narrative synthesis. As such, studies of lower quality received the same weight as studies of high quality. Additionally, although we categorized findings based on one coding system that aligned with previous research (Chang et al., 2012; Pontifex et al., Submitted for publication), other researchers may have categorized outcomes differently. As the science evolves and leads to more precise measurement, it is likely that some of our results might be coded under other headings or integrated into higher-order categories. Although cognitive function outcomes were examined in many studies, there were very few studies that examined brain function and brain structure using a RCT design, so these results should be seen as tentative and interpreted with caution until more evidence is available. Moreover, many of the results that were synthesized for brain function and brain structure came from the same subsamples of participants. Another limitation is that we only examined studies that included apparently healthy populations and therefore, our results may not generalize to other populations such as children with neurodevelopmental disorders such as Attention Deficit Hyperactivity Disorder. Additionally, grey literature was not included in our search parameters. Researchers in the future may wish to include grey literature to determine if findings are consistent across different sources of research. In an effort to provide standardized results across studies, we did not include studies that utilized non-standardized outcomes of brain health, which include indicators of grade point average or results from teacher-created tests. Researchers may wish to examine such studies to determine the generalizability of these results. During the review process, studies were brought to our attention by reviewers that were not captured by our search strategy. A decision was made not to retrospectively include nominated papers identified by anonymous reviewers as it could add bias to our review since those studies identified may only represent a subset of additional papers not captured, published, or indexed at the times of our searches. It is therefore possible that additional studies have been published that could contribute to this body of work. Supplementary File 10 contains a list of studies that were screened in full text and their reason for exclusion. In future systematic reviews, researchers could have their search strategies peer-reviewed per PRESS guidelines to ensure the robustness of the search strategy (McGowan et al., 2016).

We excluded studies that used exergames or screen-based PA exposures to minimize confounding (Stammore et al., 2017) and because they are not recommended as a public health strategy to increase overall physical activity or fitness (Chaput et al., 2013). Notwithstanding, others have argued that exergames could enhance cognitive engagement (Best, 2013). For example, some researchers have found that exergaming increased executive function in children and that was not confounded by cognitive engagement (Best, 2012). Other experimental research with youth has shown that exergames with high cognitive engagement could be beneficial for indicators of cognition compared to physical activity with low cognitive engagement or sedentary while watching a video (Benzing, Heinks, Eggenberger, & Schmidt, 2016). Nonetheless, others have argued that exergaming could be cognitively fatiguing leading to no effect (O’Leary, Pontifex, Scudder, Brown, & Hillman, 2011). In the future, researchers may wish to further explore exergaming as a PA exposure and its impact on cognitive function, brain structure, and brain function, especially with the emergence of new exergames that might contribute to overall physical activity levels (e.g., virtual reality gaming). Finally, researchers are encouraged to employ Bayesian statistics to directly test the null hypotheses to examine if acute and/or chronic PA has no effect on brain health in children and youth.

5 Conclusions

Our systematic review of RCTs on PA and brain health in apparently healthy children and youth < 18 years provides further evidence that PA is generally unrelated or beneficial for cognitive function, brain function, and brain structure. At the least, given the large number of null and favourable results for cognitive function, we suggest that PA would not impede cognitive function and should be promoted to support it.

Acknowledgements

The authors would like to thank Darcie Valois for their assistance in screening articles, Veronica Zuccala and Hana Raafat for their assistance in organizing supplementary files, Margaret Sampson, MLIS, PhD, AHIP (Children’s Hospital of Eastern Ontario) for developing the electronic search strategies, Linda Slater, MLIS (John W. Scott Health Sciences Library, University of Alberta) for peer review of the MEDLINE search strategy, and the ParticipACTION Expert Statement on Brain Health panel for the feedback on preliminary results from this study and suggestions on how to code and stratify the results. Funding for this project was provided by a grant from The Organix Foundation, Canada, The Lawson Foundation, Canada, and the Children’s Hospital of Eastern Ontario Foundation, Canada.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.mhpa.2018.11.002.

References


