Mental Health and Physical Activity 16 (2019) 66–79

Contents lists available at ScienceDirect

Mental Health and Physical Activity

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

Physical activity and depression, anxiety, and self-esteem in children and youth: An umbrella systematic review

Leila Pfaeffli Dale[a,*], Leigh Vanderloo[b,c], Sarah Moore[d,e], Guy Faulkner[a]

[a] School of Kinesiology, Lower Mall Research Station 337, 2259 Lower Mall, University of British Columbia, Vancouver, V6T 1Z4, Canada
[b] ParticipACTION, 77 Bloor St W. Suite 1205, Toronto, M5S 1M2, Canada
[c] The Hospital for Sick Children, Peter Gilgan Centre for Research and Learning, 686 Bay St., Toronto, M5G 0A4, Canada
[d] Centre for Hip Health and Mobility, Vancouver Coastal Health Research Institute, 2635 Laurel St, Vancouver, V5Z 1M9, Canada
[e] Faculty of Child, Family, and Community Studies, Douglas College, 1250 Pinetree Way, Coquitlam, V2B 7X3, Canada

A B S T R A C T

Problem: A 2011 review of reviews reported small to moderate associations between physical activity (PA) and depression, anxiety and self-esteem among children and youth (aged 5–17 years). Due to the increase in reviews examining PA and mental health outcomes in children and youth over the past decade, we conducted an umbrella review to determine the current state of the literature, including whether effects were moderated by dose and type of PA, age, sex, or severity of mental illness.
Methods: We systematically reviewed literature published from 2010 onwards from six online databases to identify and summarize findings from systematic reviews examining PA and depression, anxiety, and self-esteem outcomes in children and youth. We assessed review quality using the AMSTAR 2 critical appraisal tool.
Results: We identified 26 reviews examining depression (n = 16), anxiety (n = 2), and self-esteem (n = 14). Half of the eligible reviews were considered to be of low or critically low quality (n = 13). PA had positive mental health outcomes for children and youth, specifically for reduction in depression/depressive symptoms and improvements in physical self-concept, a self-esteem sub-domain. Little research has examined PA and anxiety. The moderator analyses reviewed revealed stronger effects in populations with clinical diagnoses (e.g. depression) and for interventions consisting of regular, supervised, group-based aerobic exercise.
Conclusions: PA appears to be an effective intervention for reducing depression/depressive symptoms and improving physical self-perceptions, although additional high-quality research and moderator analyses are needed to determine what type of PA interventions may result in better mental health outcomes for children and youth.

Mental health is a multidimensional state of well-being, with both negative indicators, such as depression or anxiety, and positive indicators, such as self-esteem and self-concept. Mental illness and the negative consequences of poor mental health among children and youth are a particular public health priority. The worldwide prevalence of mental disorders in children and adolescents (aged 6–18 years) has been estimated at 13.4% (CI 95% 11.3–15.9; k = 41; pooled sample size 87,742) (Polanczyk, Salum, Sugaya, Caye, & Rohde, 2015), and there is evidence that the prevalence of poor mental health is increasing. In Ontario, Canada for example, psychological distress, which refers to symptoms of anxiety or depression, has been rising steadily among all Ontario students in Grades 7 to 12 since it was first monitored in 2013 (Boak, Hamilton, Adlaf, Henderson, & Mann, 2018).

Mental health promotion refers to the actions taken to strengthen mental health for which all can benefit even in the presence of a mental illness (Faulkner & Duncan, 2018). Physical activity (PA) participation has been advocated as a mental health promotion approach, as well as an adjunct treatment therapy for mental illness (Faulkner & Duncan, 2018). While the physical health benefits of PA are well documented (Poitras et al., 2016; Warburton, Nicol, & Bredin, 2006), the mental health benefits of PA for children and youth are less studied. Three important outcomes that are of particular relevance for the mental health of children and youth are depression, anxiety, and self-esteem. Specific to adults, systematic reviews have linked PA with fewer symptoms of depression or anxiety in clinical (Rosenbaum, Tiedemann, Sherrington, Curtis, & Ward, 2014) and non-clinical populations (Rebar et al., 2015), with some promising evidence also emerging from a meta-analysis examining PA and depression and anxiety in children and youth (Ahn & Fedewa, 2011). Self-esteem has been identified as a buffer in the onset of mental illness in childhood (Ahn & Fedewa, 2011) and an earlier systematic review showed a positive relationship between PA and self-esteem among children and youth (Ekeland, Heian, & Hagen, 2005).

Systematic reviews are used by health care practitioners as an

* Corresponding author.
E-mail addresses: leila.dale@ubc.ca (L.P. Dale), lvanderloo@participACTION.com (L. Vanderloo), smoore43@douglascollege.ca (S. Moore), guy.faulkner@ubc.ca (G. Faulkner).

https://doi.org/10.1016/j.mhpa.2018.12.001
Received 13 September 2018; Received in revised form 29 November 2018; Accepted 6 December 2018
Available online 11 December 2018
1755-2966/

efficient way to gather evidence on a particular treatment or prevention option. There is a danger in accepting the evidence of a single systematic review, without assessing its quality (Shea et al., 2017). An umbrella review, or a review of reviews, summarizes all available evidence, compares and contrasts findings, aids in understanding the strength of the evidence, and identifies gaps in the literature (Aromataris et al., 2015). A review of reviews examining the mental health outcomes associated with PA in children and youth, in particular depressive symptoms, anxiety, self-esteem, and cognitive functioning was published in 2011 (Biddle & Asare, 2011). The strongest effects were found for self-esteem. PA showed potential benefits on reduced depression and small benefits for reduced anxiety, but both outcomes were understudied (Biddle & Asare, 2011). Many gaps regarding the effect of PA on mental health outcomes in children and youth were identified, including whether the effects of PA were strongest for those with poorer mental health at baseline (Biddle & Asare, 2011). The authors argued additional research was needed as most studies were cross-sectional, small scale, and lacked measurement consistency.

In Canada, a diverse group of paediatric researchers and practitioners in neuroscience and exercise science recently collaborated to develop an evidence-informed *Expert Statement on Physical Activity and Brain Health for Children and Youth* (under 18 years) (Vanderloo et al., 2018). The *Expert Statement* was created to highlight the positive relationship between PA and brain health, capturing both cognitive function and mental health, both within typically-developing children and youth and those with neurodevelopmental disabilities. The launch of the *Expert Statement* and the release of the *ParticipACTION Report Card on Physical Activity for Children and Youth* took place on June 19, 2018. The expert statement was largely informed by a systematic review conducted by Gunnell and colleagues (in press; registered with PROSPERO, CRD42016042116), which examined the relationship between PA and cognition, brain function, and brain structure in children with typical development (1 month–18 years).

To provide further supporting evidence underpinning the *Expert Statement*, the purpose of this umbrella review was to determine the current state of the literature, including how the field has progressed and what gaps and inconsistencies remain in terms of the relationship between PA and depression, anxiety, and self-esteem in children and youth. More specifically, we were interested in determining: 1) if new longitudinal evidence has emerged; 2) the intervention effects of different doses of PA (frequencies, intensity, duration, and types of PA); and 3) whether effects were moderated by age, sex, or severity of mental illness (e.g. prevention or treatment of depression or anxiety). We also felt it was important to assess the methodological quality of the systematic reviews included, as this was not done as part of the original review (Biddle & Asare, 2011).

1. Methods

A review protocol was created by the lead author prior to conducting the search and was agreed upon by all authors, which included the review question, search strategy, and the inclusion/exclusion criteria. Any changes in the methods that deviated from the guide are described in the present paper.

1.1. Inclusion criteria

This umbrella review was designed with broad inclusion criteria. Reviews were required to be a systematic review or meta-analysis design to ensure the highest possible level of evidence was summarized (Aromataris et al., 2015). We followed the PICO format (Liberati et al., 2009) as outlined below. A review was included if it contained at least two studies that met the inclusion criteria. If the review included other physical or mental health outcomes (e.g. bone health, academic achievement), qualitative study designs, or adult populations, the review was included if data from at least two eligible studies were able to be extracted separately from other outcomes or populations.

1.1.1. Participants

We included reviews that assessed school-aged children and youth who were apparently healthy with typical development (mean age between 5 and 17.99 years old). Reviews that focused exclusively on children and adolescents with cognitive, behavioural, or development disabilities were excluded; however, in light of our research questions, reviews with samples of children or adolescents clinically diagnosed with depression or anxiety were included. Reviews of populations with overweight/obesity were also included. If a review contained studies with both typically-developing and children or youth with disabilities, or if it included studies on adults, the review was included if the data from populations with typical development could be extracted separately or if the review included a moderator analysis on sample type (e.g. sample type did or did not moderate the effect of PA on the mental health outcome).

1.1.2. Interventions

We included reviews that reported chronic PA studies, including interventions, using quantitative measures. Reviews of acute PA (e.g. single bouts) studies were excluded unless any studies of chronic PA (e.g. repeated bouts) could be extracted from the main results. We distinguished PA from fitness and excluded reviews that examined the relationship or effects of fitness on mental health outcomes. Reviews that focused exclusively on yoga or mindfulness interventions were also excluded. Reviews that included multiple healthy behaviours (e.g. PA and diet) were included if data from studies that had PA as their main intervention component could be extracted separately from the main results.

1.1.3. Comparators

To keep the broad scope of this umbrella review, we accepted any type of comparator group for reviews of chronic PA intervention studies.

1.1.4. Outcomes

We included studies that reported the relationship between PA and at least one of the following quantitatively measured mental health outcomes.

- Depression, including depressive disorder, depressive/depression symptoms, or depressed mood.
- Anxiety, including general or social anxiety but excluding post-traumatic stress disorder.
- Self-esteem, encompassing global or physical domains of self-worth and self-concept, and the sub-domains of perceived competence and perceived physical appearance/body image (Lubans et al., 2016). General self-efficacy outcomes were excluded, as this is typically defined as confidence to overcome barriers to PA participation, not to perform PA itself (Babic et al., 2014).

A deviation from the review guide was the exclusion of reviews examining PA and cognitive function. Appraisals of cognitive functioning were synthesized in a separate review of reviews given the focus of the Expert Statement and an additional systematic review examining the effects of PA on brain function and structure (Gunnell et al., in press).

1.2. Search strategy

To retrieve eligible reviews of studies examining the relationship or effects of chronic PA and mental health in children and adolescents, the following databases were searched from inception up to December 2017: Medline, PubMed, SPORTDiscus, PsychINFO, Web of Science, and Cochrane Library. Terms that reflected the exposure variable of

Table 1
Example of the search strategy used for the Medline database.

Search No.	Search terms	No. of records found
1	(Physical activ* or exercis* or fitness or sport*).ab,ti.	441976
2	(dance or recreation or leisure or play*).ab,ti.	1098963
3	Physical education and Training	14605
4	Or/1-3	1508685
5	(depress* or despair or despondency).ab,ti.	424421
6	(anxiety or anxious or stress*).ab,ti.	871963
7	(Self-esteem or self-worth).ab,ti.	19645
8	Cognition/	90412
9	Mental health.mp or Mental Health/	165851
10	Or/5-9	1399241
11	(Child* or youth or adolescent* or teen).ab,ti.	1447811
12	(boy* or girl*).ab,ti.	230455
13	(young person or young people).ab,ti.	23906
14	Or/11-13	1569824
15	(review or narrative or summary or systematic or meta-analysis).ab,ti.	1752087
16	4 and 10 and 14 and 15	1405
	Filter by publication date: 2010–2017	884

interest (e.g. PA, exercise, sport, fitness, dance, recreation, leisure, physical education), mental health outcome variables of interest (e.g. depression, anxiety, self-esteem), population (e.g. children, adolescents, youth) and methods (e.g. systematic, narrative, review, meta-analysis) were searched. Search results were then filtered by publication date, from 2010 onwards. See Table 1 for an example of a complete electronic database search strategy.

The search and full-text review were conducted by the lead author; a second author independently reviewed 20% of the full-text articles and achieved good agreement, with all authors reaching consensus on the final number of included reviews. Any reviews included in Biddle and Asare (2011) were excluded during the full-text review, as well as reviews written in a language other than English. Reviews not published in a peer reviewed journal were excluded to ensure rigour. Citation tracking was used to examine the reference lists of included reviews to check for any potential reviews not identified during database searching.

1.3. Data extraction

All authors agreed on the variables to include in the data extraction table, based on Biddle and Asare's (2011) review and items recommended by Aromataris et al. (2015), to aid in comparing our updated results. Extracted data included the years of publication the review covered, characteristics of the population, the PA exposure variable, mental health outcome measured, types of study design, and main findings. Data was extracted by one author; a second author independently checked 20% of the studies against the data extraction tables and found no discrepancies. In addition, we classified the main findings of reviews as positive, negative, uncertain, or no association, adapting the method first developed by Sallis, Prochaska, and Taylor (2000) (Table 2).

1.4. Quality assessment

Systematic reviews summarize the best available evidence on a topic; however, the quality of the planning, conduct, and writing of a systematic review should be appraised before basing decisions on its results (Shea et al., 2017). AMSTAR 2 is a 16-point critical appraisal tool of systematic reviews of randomized controlled trials and non-randomized intervention studies (Shea et al., 2017). We used AMSTAR 2 to evaluate the quality of the systematic reviews that included at least two intervention studies (e.g. RCTs, non-RCTs, experimental and quasi-experimental study designs). The quality of systematic reviews that only included cross-sectional or longitudinal studies (i.e. no intervention studies) was not assessed. Two reviewers independently assessed the quality of each included review that met the AMSTAR 2 criteria. Any discrepancies were discussed, and consensus was reached.

1.5. Number of unique studies

We were also interested in determining the number of unique individual studies included in the reviews, to ensure that each review was not simply reviewing the same studies. To calculate the frequency counts, we kept a list of included studies in each review and counted the number of unique studies published in 2007 or later, as the reviews included in Biddle and Asare's (2011) review of reviews searched up to 2006.

2. Results

From the six databases searched, 26 distinct reviews met the inclusion criteria. Fig. 1 illustrates the search results.

In Biddle and Asare's (2011) review, four review articles were identified concerning depression, four for anxiety and three for self-esteem. Our search yielded 26 new reviews on PA and mental health outcomes, specifically depression (n = 16), anxiety (n = 2), and self-esteem (n = 14). Four reviews included studies with data on more than one relevant mental health outcome (Eime, Young, Harvey, Charity, & Payne, 2013; Evans et al., 2017; Hermens, Super, Verkooijen, & Koelen, 2017; Lubans et al., 2016). A list of excluded reviews, with the reasons for exclusion, can be found in Supplemental File 1. The data extraction for each included review can be found in Tables 3–5. Of the 26 included reviews, there were 150 unique studies published after 2006, the most recent search date of reviews included in Biddle and Asare's (2011) review, justifying the need to conduct an updated umbrella review.

2.1. Quality assessment

We classified seven reviews as moderate quality, and the remainder as low (n = 9), critically low (n = 4), or not eligible to be assessed (n = 6) (see Supplemental File 2). Items that were most frequently not met or omitted in the reviews included *item 2: establishment of review methods prior to conducting the review*; *item 3: explanation of study design selection*; *item 7: providing a list of excluded studies with justifications*; *item 10: reporting on the sources of* funding *for the included studies*. Most reviews (18/20; 90%) did use a satisfactory technique for assessing risk of bias in the included individual studies (*item 9*) and discussed (16/20;

Table 2
Classification of included systematic reviews.

Rationale	Classification
0–33.3% of studies reported a significant association, or no treatment effect observed in meta-analysis	0 = no association
34–59% of studies reported a significant association, or if less than four studies were included	? = inconsistent or uncertain association
60% of studies found a significant association, or treatment effect observed in meta-analysis	+ or - = positive or negative depending on the direction of the association

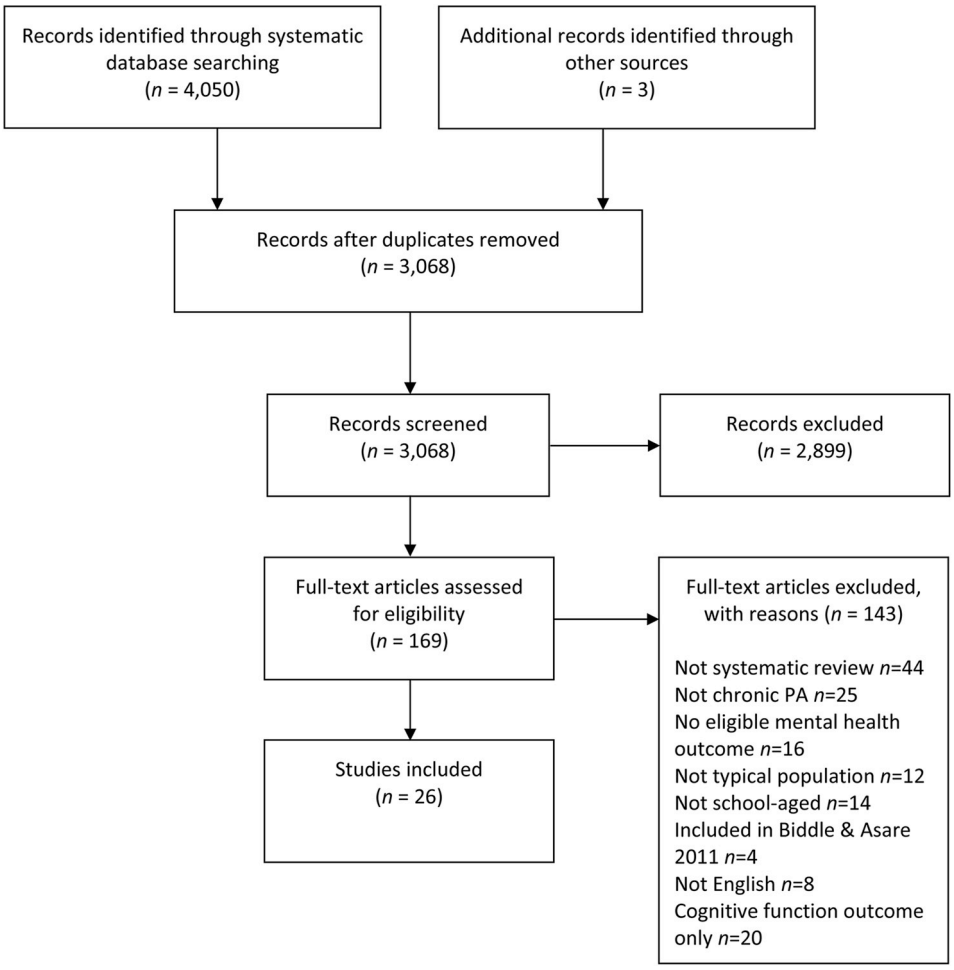

Fig. 1. Flow diagram of included studies.

80%) how bias may have affected the findings of the review (*item 13*).

2.2. Depression

Since Biddle and Asare's (2011) review, we found an additional 16 reviews of PA and depression in children and youth. Review characteristics and main findings are presented in Table 3.

2.2.1. Study characteristics

2.2.1.1. Design. Of the 16 included systematic reviews, five conducted meta-analyses and 11 conducted a qualitative summary only. The design of studies included in the reviews varied; two meta-analyses included only RCTs (Bailey, Hetrick, Rosenbaum, Purcell, & Parker, 2017; Carter, Morres, Meade, & Callaghan, 2016), four reviews included intervention studies only (Ash, Bowling, Davison, & Garcia, 2017; Brown, Pearson, Braithwaite, Brown, & Biddle, 2013; Lubans et al., 2016; Radovic, Gordon, & Melvin, 2017), four reviews included only cross-sectional, observational or longitudinal designs (Dennison, Sisson, & Morris, 2016; Eime et al., 2013; Hoare, Skouteris, Fuller-Tyszkiewicz, Millar, & Allender, 2014; Korczak, Madigan, & Colasanto, 2017), and the remainder included a combination of design types (n = 6).

2.2.1.2. Population. All the reviews included studies of typically-developing children and youth, with four reviews having more defined populations. Data extracted from three reviews included only samples of children and youth experiencing a clinical diagnosis or threshold symptoms of depression (Ash et al., 2017; Bailey et al., 2017; Radovic et al., 2017). Hermens et al. (2017) only included studies with socially vulnerable youth (i.e. youth living in areas of low socioeconomic status, living in residential care, or receiving nonresidential counseling).

2.2.1.3. Variables. For the reviews that included experimental studies, most PA interventions consisted of supervised, aerobic, and group-based exercise. Four reviews (Eime et al., 2013; Evans et al., 2017; Hermens et al., 2017; Johnson & Taliaferro, 2011) specifically examined sports participation. When comparison groups were mentioned, they mostly consisted of a combination of no treatment, wait-list, or treatment as usual (Bailey et al., 2017; Carter et al., 2016). Some reviews reported active comparison groups, such as groups engaged in flexibility exercises (Janssen & Leblanc, 2010) or less sport (Eime et al., 2013). Most reviews included self-report measures of PA, with one review including studies that used objective measures of PA only (Poitras et al., 2016).

Regarding the outcome variable of depression, some reviews included studies with depression classification scores, such as major depressive disorder, or with continuous scales of depressive symptoms or depressed mood, or both. In this umbrella review, we used the general term "depression," which is comprised of all of the depression measures from each included review (e.g. depression, major depressive disorder, depressive/depression symptoms, depressed mood). We specified the outcomes where possible in the findings and Table 3.

2.2.1.4. Overall findings. Overall, the reviews demonstrated PA was inversely associated with depression. Twelve reviews (12/16; 75%)

Table 3
Reviews of physical activity and depression in children and youth.

Author, date and years covered	Type of review; number of studies (K)	Sample for current analyses	Exposure variables (PA) [comparison groups if specified]	Outcome variables (mental health)	Types of research design	Main findings	Classification
Ash et al. (2017); Up to June 2016	Systematic review K = 24 (K = 2 for children diagnosed with depression; remainder were participants diagnosed with ADHD/autism/bipolar)	< 21 years (diagnosed with social, emotional, or behavioural disability)	Quantitative analysis of PA intervention; aerobic; repeated exposure (6–12 weeks)	Behavioural, psychological, or cognitive outcomes	RCTs, crossover studies	• 1/2 studies had positive findings post-intervention. • 2/2 studies had positive findings for depression at 6-month follow up; 1 of these had NS findings at 12 m follow-up	?
Bailey et al. (2017); 1980–2016	Meta-analysis K = 16	12–25 years (experiencing a diagnosis or threshold symptoms of depression)	Any PA intervention; most trials were supervised MVPA; 5–12 weeks [control groups = no-treatment, wait-list, attention/activity placebo (e.g. stretching)]	Depression symptoms assessed with validated symptom scale	RCTs	• Large treatment effect compared to controls (SMD = − 0.82, 95% CI = − 1.02 to −0.61, $p < 0.05$, $I^2 = 38\%$). • Effect sizes did not differ by type of control group, age group (under or over 18 yrs), diagnostics status, depression severity category, type of PA, and intensity of PA. • PA may produce a clinically significant reduction in depression symptoms.	-
Brown et al. (2013); Up to May 2011	Meta-analysis K = 9	5–19 years	Interventions to promote or increase PA; supervised aerobic exercise most common; 9–40 wks	Quantitative measure of depression	RCTs, controlled trials, quasi-experimental	• Small summary treatment effect for depression (Hedges' g = −0.26, SE = 0.09, 95% CI = −0.43, −0.08, p = 0.004). • Methodological (e.g. studies with both education and PA intervention; those with a higher quality score; and less than 3 months in duration) and participant characteristics (e.g. single-gender studies; those targeting overweight or obese groups) contributed most to the reduction in depression.	-
Bursnall (2014); 2008–2013	Systematic review K = 10	11–17 years	Measure of PA The only RCT was supervised, regular VPA	Change in depressive symptoms	Nonexperimental cohort, RCT, quasi-experimental	• Consistent strong inverse correlation between PA and depressive symptoms (10/10 studies).	-
Carter et al. (2016); Up to April 2014	Systematic review & Meta-analysis K = 11; (8 trials eligible for meta-analysis)	13–17 years	Interventions that promoted exercise or PA; supervised group exercise; 6–40 wks; majority aerobic exercise [comparison: no treatment, wait-list, treatment as usual, psychosocial/educational intervention]	Validated outcome measure of depressive symptoms	RCTs	• Moderate overall treatment effect on depressive symptom reduction (SMD = −0.48, 95% CI = -0.87, 0.10, p = 0.01, $I^2 = 67\%$). • Trials with exclusively clinical samples showed a moderate effect on depressive symptoms with lower levels of heterogeneity (SMD = −0.43, 95% CI = −0.84, −0.02, p = 0.04, $I^2 = 44\%$). Trials without clinical samples recorded a NS moderate effect.	-
Dennison et al. (2016), p. 1990–August 2015	Systematic review K = 21; (K = 15 for PA; remainder assessed other outcomes)	6–12 years	At least 1 measure of PA, most used valid/reliable self-report measure of PA	Measure of depression or depressive symptoms	Cross-sectional or longitudinal cohort (interventions excluded)	• 66% of studies reported PA was inversely related to depressive symptoms. • Age and gender may be mediating factors; PA may have a greater effect on depressive symptoms in younger females and older males.	-
Eime et al. (2013), p. 1990–May 2012	Systematic review K = 30 (K = 4 for depression; remainder measured other outcomes or were qualitative)	Children and adolescents (no ages given)	Sport [majority of studies compared sport to no sport or other extra-curricular activities]	Mental and/or social health	Cross-sectional, observational, or longitudinal	• 4/4 studies (3 cross-sectional, 1 longitudinal) showed fewer depressive symptoms/lower depression scores associated with sport participation compared to less or no sport participation.	-
Evans et al. (2017), p. 1980 to May 2016	Systematic review K = 35 (K = 4 for depression; remainder assessed other outcomes)	7–17 years (youth sport participants)	Sport activity constructs [team vs individual sport participation]	Psychosocial constructs	Quantitative studies that contrasted activity types or levels of activity	• 2/2 studies showed more frequent team sport participation was negatively associated with depressive symptoms than no sport.	?

(*continued on next page*)

Table 3 (*continued*)

Author, date and years covered	Type of review; number of studies (K)	Sample for current analyses	Exposure variables (PA) [comparison groups if specified]	Outcome variables (mental health)	Types of research design	Main findings	Classification
						• 1/2 studies showed fewer depressive symptoms for youth who took part in team vs individual sport.	
Hermens et al. (2017); 1990–2014	Systematic review K = 18 (K = 2 for depression; remainder assessed other outcomes or were qualitative)	10–23 years (socially vulnerable youth, low SES or living in residential/non-residential care)	Sports program, majority in schools	Life skills (includes anxiety, depression, & self-esteem)	Experimental quantitative to qualitative interviews	• 1/2 studies found reduced depressive symptoms for those who participated in sport.	?
Hoare et al. (2014); 2002–2013	Systematic review K = 24 (K = 12 for PA; remainder assessed other outcomes)	10–19 years (typically developing)	Obesogenic risk factors	Measure of depression	Cross-sectional or longitudinal cohort	• 4/6 cross-sectional studies found inverse relationships between depression on PA: 2 studies showed inverse relationship with PA. 2 studies showed involvement in sporting clubs or extracurricular activities were associated with less depressive symptoms. • 4/6 longitudinal studies found PA was inversely related to depressive outcomes over a 1–6 year period.	-
Janssen and Leblanc (2010); Up to 2008	Systematic review K = 86 (K = 6 for depression; remainder assessed other outcomes)	5–18 years	Level of PA; RCTs were prescribed aerobic exercise, 60–90 min/wk	Depression (as a measure of mental health)	x-sectional; RCT	• Small-modest treatment effects were found in all 3 RCTs. Relationship more evident for moderate PA than VPA. • 1/3 cross-sectional studies reported a modest relationship between PA and depression.	?
Johnson and Taliaferro (2011); 1997–2010	Systematic review K = 19	14–19 years (mean age)	MVPA and sports team participation	Measure of depressive symptoms	Cross-sectional, longitudinal, intervention	• 18/19 studies found inverse relationships between PA, particularly sports participation, and depressive symptoms.	-
Korczak et al. (2017); 2005–2015	Meta-analysis K = 40	< 18 years	Measure of PA	Validated measure of depressive symptoms	Cross-sectional or longitudinal	• Association between PA and depression was found ($r = -0.14$; 95% CI = −0.19 to −0.10). Stronger effect found in studies with: • cross-sectional vs longitudinal designs • using depressive symptom self-report vs diagnostic interview • using validated vs non-validated PA measures • using measures of frequency and intensity of PA versus intensity alone	-
Lubans et al. (2016); Up to July 2015	Systematic review K = 25 (K = 4 for depression; remainder assessed other outcomes)	5–18 years	Any type of school/home/community PA or laboratory-based exercise intervention	Changes in mental health outcomes (self-concept/self-esteem & depression)	RCTs, quasi-experimental	• 3/4 studies found negative effects of PA on depression post-intervention.	-
Poitras et al. (2016); Up to January 2015	Systematic review K = 162 (K = 4 for depression; remainder measured other outcomes)	5–17 years (apparently healthy)	Objectively measured PA	Health indicators (includes depressed mood, depressive symptoms, major depressive disorder)	All designs considered, but had to reach minimum sample size (30 for RCTs; 1000 for x-sectional)	• Cross sectional studies: Some support for favourable associations between total PA & MVPA and depression o Total PA: favourable in 1/4 studies o VPA: NS (1 study) o MVPA: favourable in 1/4 studies o LPA: NS (1 study) • Longitudinal study: no association between baseline MVPA and depressed mood at follow up (1 study)	0
Radovic et al. (2017); Up to January 2015	Meta-analysis K = 8	12–18 years (clinical levels of depressive symptoms)	Exercise was primary or significant component of intervention (2x/wk for > 20min); majority used	Measure of depressive symptoms	RCTs or controlled trials	• Moderate reduction in depressive symptoms post-intervention (Hedge's $g = -0.61$, $p = 0.007$)	-

(*continued on next page*)

Table 3 (*continued*)

Author, date and years covered	Type of review; number of studies (K)	Sample for current analyses	Exposure variables (PA) [comparison groups if specified]	Outcome variables (mental health)	Types of research design	Main findings	Classification
			group-supervised aerobic exercise, 4–20wks			● Dose of exercise did not significantly predict reduction in depressive symptoms	

Note: PA = physical activity; RCT = randomized controlled trial; NS = non-statistically significant; MVPA = moderate-to vigorous-intensity physical activity; VPA = vigorous physical activity; LPA = light physical activity; ? = uncertain association; - = negative association; 0 = no association.

were classified as showing a negative association, meaning that higher levels of PA were associated with lower levels of depression. We classified three reviews as uncertain (3/16; 19%), due to the small number of studies in each review (range 1–3). One review was classified as showing no association (1/16; 6%); interestingly, this was the sole review to include studies with only objective measures of PA (Poitras et al., 2016). All five meta-analyses showed PA reduced depressive symptoms. Specifically, PA interventions had small (Brown et al., 2013) to moderate (Carter et al., 2016) treatment effects on depressive symptoms in children and youth with typical development. Two reviews that included only populations with clinically diagnosed depression found moderate (Radovic et al., 2017) to large (Bailey et al., 2017) treatment effects. The meta-analysis examining cross-sectional studies found small consistent relationships between PA and depression/depressive symptoms (Korczak et al., 2017). The same review also found a weak effect size for longitudinal associations between PA and depressive symptoms, measured 2–17 years later (Korczak et al., 2017). Three (Dennison et al., 2016; Hoare et al., 2014; Korczak et al., 2017) of four (Poitras et al., 2016) reviews with longitudinal studies showed PA protected against depression.

2.2.1.5. Moderators. Moderator analyses yielded mixed results. One review concluded that PA may be more effective for populations with clinical diagnoses (e.g. depression) compared to populations with typical development (Carter et al., 2016), which corresponded with the largest overall effect size found by Bailey et al. (2017), who only included children and youth diagnosed with depression or experiencing threshold level symptoms. Bailey et al. (2017) found that their effect size did not differ by diagnostic status (threshold vs. diagnosed) or depression severity.

One review found greater treatment effects for higher quality studies (clinical trials/RCTs) (Brown et al., 2013), which is important as many reviews reported low quality of studies (Bailey et al., 2017; Carter et al., 2016; Hoare et al., 2014; Radovic et al., 2017). Korczak et al. (2017) found stronger effects in cross-sectional over longitudinal designs, as well as in those studies using self-report of depressive symptoms compared to gold-standard diagnostic interview to diagnose depression.

There were mixed findings in terms of the most effective type of PA intervention; however, supervised, aerobic-based group exercise, engaged in multiple times per week over at least 7–8 weeks appeared to be most effective in reducing depression (Bailey et al., 2017; Carter et al., 2016). Two reviews on populations with clinical depression found the intensity (Bailey et al., 2017) or dose (Radovic et al., 2017) of PA had no effect on the reduction of depressive symptoms. For sport-only reviews, the level of sport involvement (intensity or frequency) was not described (Eime et al., 2013; Evans et al., 2017; Hermens et al., 2017). It was not possible to summarize the moderating effects of age (Bailey et al., 2017; Dennison et al., 2016), sex, or weight status (overweight/obese vs average weight) (Brown et al., 2013) as inconsistent effects were found.

2.3. Anxiety

Since Biddle and Asare's (2011) review, two new systematic reviews met our criteria for PA and anxiety in children and youth (Evans et al., 2017; Hermens et al., 2017). Both reviews combined included a total of five unique studies with anxiety outcomes (Table 4).

2.3.1. Study characteristics

2.3.1.1. Design. The design of the included studies in each review varied, with one review including only cross-sectional, observational, or longitudinal studies (Eime et al., 2013) and the other included all design types, from qualitative to experimental (Hermens et al., 2017).

2.3.1.2. Population. One review included studies on children and

Table 4
Reviews of physical activity and anxiety in children and youth.

Author, date and years covered	Type of review; number of studies (K)	Sample for current analyses	Exposure variables (PA)	Outcome variables (mental health)	Types of research design	Main findings	Classification
Eime et al. (2013), p. 1990–May 2012	Systematic review K = 30 (K = 3 for anxiety; remainder measured other outcomes or were qualitative)	Children and adolescents (no ages given)	Sport [majority of studies compared sport to no sport or other extra-curricular activities]	Mental and/or social health	Cross-sectional, observational, or longitudinal	• 3/3 studies (2 longitudinal, 1 cross-sectional) showed reduced anxiety was associated with sport participation • 1 of these showed a relationship for team sport only, not individual sport.	?
Hermens et al. (2017); 1990–2014	Systematic review K = 18 (K = 2 for anxiety; remainder assessed other outcomes or were qualitative)	10–23 years (socially vulnerable youth, low SES or living in residential/non-residential care)	Sports program, majority in schools	Life skills (includes anxiety, depression, & self-esteem)	All types considered; experimental, quantitative, qualitative	• 2/2 studies found reduced anxiety for those who participated in sport	?

Note: PA = physical activity; SES = socioeconomic status; ? = uncertain association.

youth, with few describing features (Eime et al., 2013), while the other review focused on socially vulnerable youth (Hermens et al., 2017).

2.3.1.3. Variables. Both reviews included sport-only studies, excluding other types of PA, and both included quantitative measures of anxiety as part of a range of mental health outcomes (Eime et al., 2013; Hermens et al., 2017).

2.3.1.4. Overall findings. Both reviews found a relationship between increased sport participation and reduced anxiety; however, they were classified as unclear associations due to the small number of studies included (Eime et al., 2013; Hermens et al., 2017). There was insufficient detail on the intensity or frequency of sport participation (Eime et al., 2013; Hermens et al., 2017), although one longitudinal study in Eime et al. (2013) found reduced anxiety over time in children and youth taking part in team sport, while those who took part in individual sport or did no sport showed increased anxiety over time.

2.4. Self-esteem

Fourteen reviews on PA and self-esteem in children and youth were included. The updated findings of reviews examining the relationship or effects of PA and self-esteem are provided in Table 5.

2.4.1. Study characteristics

2.4.1.1. Design. Two reviews limited their inclusion criteria to RCTs only (Ruotsalainen, Kyngas, Tammelin, & Kaariainen, 2015; ten Hoor et al., 2017), whereas five reviews included experimental and/or quasi-experimental studies only (Liu, Wu, & Ming, 2015; Spruit, Rafferty, (Lubans et al., 2016; Lubans, Plotnikoff, & Lubans, 2012). One review examined only associations from cross-sectional, observational, or longitudinal studies (Eime et al., 2013). The remaining six reviews included a combination of cross-sectional, longitudinal, and experimental designs.

2.4.1.2. Population. Most of the reviews included studies on children and adolescents with typical development or not-specified (Babic et al., 2014; Burkhardt & Brennan, 2012; Eime et al., 2013; Joronen, Aikasalo, & Suvitie, 2017; Liu et al., 2015; Lubans et al., 2016; Rafferty, Breslin, Brennan, & Hassan, 2016; Spruit, Assink, van Vugt, van der Put, & Stams, 2016). One review included studies on youth sport participants only (Evans et al., 2017). Two reviews limited their inclusion criteria to socially vulnerable (Hermens et al., 2017) or 'at-risk' youth who live in a negative environment and/or do not have the skills and values to help them become responsible members of society. (Lubans et al., 2012). Two reviews included only children or youth who were overweight/obese (Ruotsalainen et al., 2015; ten Hoor et al., 2017). The final review included studies examining the effects of PA on the self-esteem of boys only (Bassett-Gunter, McEwan, & Kamarhie, 2017).

2.4.1.3. Variables. Nearly half of the reviews included any type of PA intervention or measure (Bassett-Gunter et al., 2017; Liu et al., 2015; Lubans et al., 2012, 2016; Ruotsalainen et al., 2015; Spruit et al., 2016). One review included studies that measured leisure-time PA only (Babic et al., 2014). Three reviews included sport-only studies (Eime et al., 2013; Evans et al., 2017; Hermens et al., 2017), while one review each included only strength training interventions (ten Hoor et al., 2017), recreational dance interventions (Burkhardt & Brennan, 2012), exergaming interventions (Joronen et al., 2017), or specifically school-based PA interventions (Rafferty et al., 2016). Most reviews included studies with a quantitative measure of global self-esteem or self-worth, while one review exclusively focused on physical self-concept (Babic et al., 2014) and another on body image only (Bassett-Gunter et al., 2017).

2.4.1.4. Overall findings. Overall, evidence supported a small-to-

Table 5
Reviews of physical activity and self-esteem in children and youth.

Author, date and years covered	Type of review; number of studies (K)	Sample for current analyses	Exposure variables (PA)	Outcome variables (mental health)	Types of research design	Main findings	Classification
Babic et al. (2014); Up to August 2013	Systematic review & meta-analysis K = 64	4–20 years	Quantitative assessment of leisure time PA	Quantitative assessment of physical self-concept or sub-domains	Cross-sectional (n = 47), longitudinal (n = 12), and experimental (n = 5)	• A weak to moderate effect was found b/w PA and: • physical self-concept (r = 0.25, 95% CI 0.16–0.34, p < 0.001) • perceived competence (r = 0.33, 95% CI 0.27–0.39, p < 0.001) • perceived physical appearance (r = 0.12, 95% CI 0.08–0.16, p < 0.001). • Qualitative summary found strong evidence of + associations between all 4 measures and PA • Sex was a moderator for general physical self-concept (stronger association for boys) • Age was a moderator for perceived appearance (stronger for younger/early adolescents) • Age was also a moderator for perceived competence (stronger for early and late adolescents compared to children) • Study design was not a moderator	+
Bassett-Gunter et al. (2017); Up to January 2015	Meta-analysis K = 36 (K = 13 for children and adolescents; remainder assessed adult populations)	All ages (boys and men only)	Measure of PA	Measure of body image	Cross-sectional, experimental, RCT	• For adolescent males, there were significant relationships between PA and body image in correlational studies (g = 0.47, k = 9), but NS in intervention studies (g = 0.04, k = 3). • Moderator analyses (age) between exercise and body image were not conclusive	+
Burkhardt and Brennan (2012); Up to 2009	Systematic review K = 14 (K = 3 for self-worth/self-concept; remainder assessed other outcomes)	5–21 years	Recreational dance interventions [PE control]	Well-being outcomes (anxiety, physical self-worth, self-concept)	Controlled trials, cohort studies, case control studies, x-sectional with control group	• 2/3 studies showed treatment effect for physical self-worth/self-concept.	?
Eime et al. (2013), p. 1990–May 2012	Systematic review K = 30 (K = 7 for self-esteem; remainder measured other outcomes or were qualitative)	Children and adolescents (no ages given)	Sport [majority of studies compared sport to no sport or other extra-curricular activities]	Mental and/or social health (multiple)	Cross-sectional, observational, or longitudinal	• Positive findings for all 7 studies: Sport participation was associated with self-esteem (4/4), self-concept (2/2), self-image (1/1). • 4 studies were cross-sectional, 3 were longitudinal	+
Evans et al. (2017), p. 1980 to May 2016	Systematic review K = 35 (K = 3 for self-esteem; remainder assessed other outcomes)	7–17 years (youth sport participants)	Sport activity constructs [team vs individual sport participation]	Psychosocial constructs	Quantitative studies that contrasted activity types or levels of activity	• 2 studies showed more frequent team sport participation was positively associated with self-esteem than no sport. • No difference in body dissatisfaction was found between team and individual sport (1 study).	?
Hermens et al. (2017); 1990–2014	Systematic review K = 18 (K = 5 for self-esteem/self-worth; remainder assessed other outcomes or were qualitative)	10–23 years (socially vulnerable youth, low SES or living in residential/non-residential care)	Sports program, majority in schools	Life skills (includes anxiety, depression, & self-esteem)	Experimental quantitative to qualitative interviews (only quantitative included here)	• 3/3 studies found improvements in self-esteem for those who participated in sport. • 1/2 studies found improvements in global self-worth.	+
Joronen et al. (2017), p. 2004–April 2015	Systematic review K = 10 (K = 2 studies for self-esteem/self-concept; remainder assessed other outcomes)	6–18 years (non-clinical children and adolescents)	Exergaming or other active video games	Non-physical effects on well-being (includes self-esteem, perceived competence to exercise)	All design types: experimental, observational, qualitative	• Of 2 studies, 1 study showed a treatment effect of exergaming on perceived competence to exercise regularly; 1 study found no change in self-esteem.	?
Liu et al. (2015); Up to July 2014	Meta-analysis K = 38 (K = 19 studies for self-esteem; 7 studies for self-concept; 12 studies for self-worth)	3–20 years (6 studies had non-typical populations: 4 = cerebral palsy, 2 = learning or cognitively disabled)	Supervised PA intervention or PA combined with other strategies	Measure of self-esteem, self-concept, or self-worth	RCT and non-RCT studies	• Small treatment effects were found in RCTs for intervention of PA alone on: o general self outcomes (Hedges' g = 0.29, 95% CI = 0.14 to 0.45; p = 0.001)	+

(*continued on next page*)

Table 5 (*continued*)

Author, date and years covered	Type of review; number of studies (K)	Sample for current analyses	Exposure variables (PA)	Outcome variables (mental health)	Types of research design	Main findings	Classification
						o self-concept (Hedges' g = 0.49, 95% CI: 0.10 to 0.88, p = 0.014) o self-worth (Hedges' g = 0.31, 95%CI: 0.13 to 0.49, p = 0.005). o NS pooled effect size was found on self-esteem. ● NS effect of intervention of PA combined with other strategies ● Stronger effects were found in school/ gymnasium settings vs clinic, family, or detention facility settings ● Effects not altered by PA intensity, frequency, length of intervention, participant type, or study quality ● Very low heterogeneity in RCTs evaluating intervention of PA alone; lack of publication bias.	
Lubans et al. (2012); Up to December 2010	Systematic review K = 15 (K = 10 studies for self-esteem; remainder measured other outcomes)	4–18 years (disaffected or 'at-risk' youth)	PA intervention (outdoor ed, exercise, sport)	Quantitative assessment of social/emotional well-being	RCT, quasi-experimental, single group pre-post test	● Overall 6/10 studies showed treatment effect on some type of self-esteem measure: o self-worth (1/3 studies) o self-concept (2/4 studies) o self-esteem (1/2 studies) o perceived physical self-perceptions/ competence 3/4 studies). o PA had NS effect on physical attractiveness (1 study)	+
Lubans et al. (2016); Up to July 2015	Systematic review K = 25 (K = 12 for self-esteem; remainder assessed other outcomes)	5–18 years	Any type of school/ home/community PA or laboratory-based exercise intervention	Changes in mental health outcomes (self-concept/ self-esteem)	RCTs, quasi-experimental	● 7/12 studies found a + treatment effect of PA self-concept/esteem ● The effect of PA on psychosocial mechanisms of mental health: o Physical self-concept/self-worth (3/5 studies) o Perceived physical/sport/athletic competence (4/8 studies) o Perceived appearance/attractiveness or body image (7/10 studies) ● Evidence supported causal link between physical self-perceptions and self-esteem in the majority of studies	?
Rafferty et al. (2016); Up to January 2016	Systematic review K = 11 (K = 7 for self-esteem; remainder measured other outcomes)	6–12 years	School-based PA intervention	Well-being measured pre-post intervention	Experimental/quasi-experimental	● Only 1/7 studies found a positive intervention effect on self-esteem, specifically: o global self-worth (1/2 studies) o NS for body image (2 studies) o NS for self-esteem (2 studies) o NS for perceived physical self-competence/ athletic competence (4 studies) ● Only school-based PA interventions included; 5/7 studies showed some type of sig improvements in PA. Of these 5, none had sig effects on self-esteem measures	0
Ruotsalainen et al. (2015); 1950–2013	Systematic review K = 14 (K = 2 for self-esteem; remainder assessed other outcomes)	12–18 years (obese adolescents)	PA/exercise intervention (could include counseling or dietary practices)	Psychological symptoms (physical self-perception, body satisfaction)	RCTs	● 2/2 studies found treatment effect for physical self-perception/body satisfaction, but the same 2 studies found NS differences on self-esteem.	?

(*continued on next page*)

Table 5 (*continued*)

Author, date and years covered	Type of review; number of studies (K)	Sample for current analyses	Exposure variables (PA)	Outcome variables (mental health)	Types of research design	Main findings	Classification
Spruit et al. (2016); Up to August 2015	Meta-analysis K = 57 (K = 33 for self-concept; remainder measured other outcomes)	11–18 years (mean age)	PA intervention (sport or aerobic exercise)	Psychosocial outcomes (self-concept, self-worth)	Experimental	• PA interventions had a small to medium (d = 0.297) effect on self-concept • None of the sample characteristics (e.g. clinical, obese, or mental disability) moderated the effect of PA interventions on self-concept. • Larger effects on self-concept were found when the intervention consisted of (aerobic) exercise vs sport. • No moderating effects were found for the type of intervention, the duration and frequency of the intervention. • No moderating effects were found on the design of the study, the type of comparison group, and the follow-up time.	+
ten Hoor et al. (2017); Up to May 2014	Meta-analysis K = 17 (K = 2 for school aged and self-esteem; remainder assessed adult participants)	All ages (2 studies included children, ages 8–11, and youth, ages 13–17 who were overweight/obese)	Strength training interventions	Psychological outcomes	RCTs	• 1/1 study showed treatment effect on self-esteem. NS effect found on physical self-worth. • 1/1 study showed treatment effect on physical self-concept.	?

Note: PA = physical activity; RCT = randomized controlled trial; NS = non-statistically significant; PE = physical education; SES = socioeconomic status; ? = uncertain association; + = positive association; 0 = no association.

moderate relationship between increased PA and enhanced self-esteem, with half of the reviews classified as showing positive associations. One review was classified as non-associated (Rafferty et al., 2016). Six were classified as uncertain: five reviews included less than four studies (Burkhardt & Brennan, 2012; Evans et al., 2017; Joronen et al., 2017; Ruotsalainen et al., 2015; ten Hoor et al., 2017) and one review found mixed results (Lubans et al., 2016).

The included meta-analyses found PA had small associations on physical self-concept (Babic et al., 2014), small-to-moderate associations on perceived physical appearance/body image (Babic et al., 2014; Bassett-Gunter et al., 2017) and moderate associations on perceived competence (Babic et al., 2014). For meta-analyses of interventions, two reviews found small to moderate treatment effects of PA on self-concept (Liu et al., 2015; Spruit et al., 2016), while one review found small treatment effects of PA on self-worth (Liu et al., 2015). The same review found no significant pooled effect on self-esteem (Liu et al., 2015). One review found PA had no treatment effect on body image in adolescent boys; however, the authors caution that only three intervention studies were included in the analysis (Bassett-Gunter et al., 2017). Two reviews found positive associations longitudinally between PA and higher self-esteem (Eime et al., 2013) and physical self-concept (Babic et al., 2014). The meta-analysis conducted by Babic et al. (2014) showed the effect size estimates between PA and all self-concept domains were similar between experimental, cross-sectional, and longitudinal study designs.

2.4.1.5. Moderators. Patterns among the systematic reviews were difficult to identify as there were many types of PA and PA interventions examined. Two reviews found the effects of PA were not altered by PA intensity, frequency, or length of intervention (Liu et al., 2015; Spruit et al., 2016). Spruit et al. (2016) found larger effects when the intervention consisted of aerobic exercise vs sport. Liu et al. (2015) found stronger effects with school/gymnasium settings compared to clinic, family, or detention facility-based settings. In contrast, the one review classified as non-associated, showing PA had no effect on self-esteem outcomes, was a review limited to school-based PA interventions only (Rafferty et al., 2016). Study design did not alter the effects of PA in three reviews (Babic et al., 2014; Liu et al., 2015; Spruit et al., 2016).

Sample type moderators were inconsistent. Babic et al. (2014) found sex and age were significant moderators. Specifically, the authors found stronger associations between PA and physical self-concept among boys than girls, and stronger associations between PA and perceived appearance and perceived competence among early adolescents compared to children. Spruit et al. (2016) found the PA interventions were equally effective for children and youth of healthy weight as well as overweight/obese. Two reviews found no moderating effects of participant type (typically-developing vs. children/youth with disabilities) (Liu et al., 2015; Spruit et al., 2016).

3. Discussion

We found 26 reviews examining PA and the mental health outcomes of depression (n = 16), anxiety (n = 2), and self-esteem (n = 14) in children and youth. In the case of depression and self-esteem, the number of systematic reviews has quadrupled in eight years. To a certain extent this reflects a rapidly maturing body of evidence with 150 new studies being assessed within those reviews. Since the Biddle and Asare (2011) review, evidence supporting the associations between PA and reduced depression or depressive symptoms, and improvements in physical self-perceptions, has grown. Concerns that the evidence for these associations is limited should be tempered; however, research examining PA and anxiety has stagnated. The findings for each mental health outcome will be discussed separately, followed by an overall discussion of strengths and limitations.

3.1. Depression

Our umbrella review demonstrated that PA was inversely associated with depression, with 12/17 (71%) reviews showing increased PA was associated with or led to decreased depression or depressive symptoms. The inclusion of five meta-analyses with small-to-moderate treatment effects strengthens the evidence behind causality, although most reviews concluded there were potential risks of bias. Larger effects of PA were found for children and adolescents clinically diagnosed with depression, with one review concluding that PA may produce a clinically significant reduction in depressive symptoms (Bailey et al., 2017). Similarly, exercise has also been shown to reduce depression in adults (Schuch et al., 2017) and older adults (Catalan-Matamoros, Gomez-Conesa, Stubbs, & Vancampfort, 2016). Meta-analyses of adult populations also saw larger effect sizes of PA on depression for populations with clinical depression (Rosenbaum et al., 2014) compared to non-clinical (Rebar et al., 2015).

The contrasting finding of the one review classified as showing no relationship between PA and depression included studies that only used objective measures of PA (Poitras et al., 2016), while all other reviews included self-report. Studies in other reviews using self-report measures of PA may have found greater effects due to social desirability and could have risks of detection bias (Bailey et al., 2017). Notably, Poitras et al. (2016) only reported results from four cross-sectional and one longitudinal study, which limits their findings. Future research should consider objective measures of PA; meta-analyses should conduct moderator analyses to determine if type of PA measure (self-report or objective) leads to different effects on depression. One review that did look at the moderating effect of measurement type, albeit for depression not PA, found larger effects for studies using self-report measures of depressive symptoms compared to the gold-standard diagnostic interview (Korczak et al., 2017). Only a small number of studies used the interview assessment ($n = 4$, compared to $n = 46$ for self-report), which also questions the strength of the current evidence. From the evidence in this umbrella review it appears that PA was linked with reduced depressive symptoms. It was difficult to determine if PA specifically affected depression diagnosis, as many reviews included studies measuring different presentations of depression and used the terms depression and depressive symptoms interchangeably. Future studies should consider using clinical interview to assess depression and be specific with terminology.

We found inconsistent moderator analyses of PA dose and type; however, interventions that included moderate-intensity, supervised aerobic and resistance training appeared to be most effective (Bailey et al., 2017; Carter et al., 2016; Janssen & Leblanc, 2010). One review found even a modest amount of prescribed exercise (60–90 min per week) had a small treatment effect on depression (Janssen & Leblanc, 2010), while another review found the dose of exercise did not predict reduction in depressive symptoms (Radovic et al., 2017). More dose-response trials are needed before specific treatment guidelines for depression can be made. For now, following the commonly recommended national PA guidelines for children and youth (5–17 years) of 60 minutes of moderate-to vigorous-intensity PA per day and muscle and bone strengthening activities three days per week (Tremblay et al., 2016) may result in the greatest overall mental and physical health benefits. Our review found evidence that lower doses of PA were still associated with mental health benefits, thus simply ‘moving more’ is likely a good starting message, as the most effective dose is likely the one that can be maintained.

3.2. Anxiety

Our findings showed a relationship between increased PA and reduced anxiety; however, the evidence was limited to two reviews, and we classified the overall relationship as unclear. Biddle and Asare (2011) found a small-to-moderate effect on anxiety, summarized from four reviews with individual studies published up to 2006. Since then, there has been little advancement in the evidence base regarding PA and anxiety in children and youth. A systematic review should be conducted that specifically examines the effects of PA on anxiety in children and youth from 2006 onwards in individual studies. Considering the wider literature, existing evidence suggests that PA interventions may have a broadly reductive effect on anxiety symptoms, but this remains a relatively novel focus of research (Kandola et al., 2018).

One potentially informative review was excluded from this umbrella review because studies with children/youth with disabilities and those with typical development were considered (Ahn & Fedewa, 2011). The authors found PA was associated with reductions in anxiety, measured in 16 studies, although a list of included studies was not provided, thus we were unable to tease out results from individual studies meeting our criteria. The authors did not conduct moderator analyses specific to differences in depression, anxiety, or self-esteem between population groups. The authors’ findings did match some of the findings from other included reviews, particularly that PA reduced depression and anxiety and improved self-esteem in children and youth, and that PA was more effective at improving mental health in clinical than non-clinical populations (Ahn & Fedewa, 2011).

3.3. Self-esteem

PA had a weak-to-moderate effect on self-esteem outcomes in children and youth, and half of the reviews (7/14; 50%) showed positive findings. Only one study reported no intervention effects on self-esteem outcomes (Rafferty et al., 2016), but five reviews were classified as uncertain due to the small number of included studies. In comparison, Biddle and Asare's review (2011) also found small-to-moderate associations or effect of PA on self-esteem. The most consistent evidence was for self-concept, with three meta-analyses showing treatment effects of PA on improved physical self-concept (Babic et al., 2014; Liu et al., 2015; Spruit et al., 2016). The findings should be interpreted with caution; however, as many reviews rated individual study quality as poor with high risk of bias (Burkhardt & Brennan, 2012; Lubans et al., 2012; Rafferty et al., 2016; Spruit et al., 2016). More research with larger samples, using robust methods (i.e. randomized controlled trials with inactive comparison groups) are needed to advance the field and truly examine causality.

Moderator analyses from included reviews revealed PA positively influenced self-esteem outcomes in all types of populations, both typically-developing, overweight/obese, and in children and youth with disabilities. As with depression, important questions remain on the type and dose of PA intervention that are most effective. The PA interventions were varied, but there was some agreement that supervised, aerobic-based interventions were most effective.

One explanation for the conflicting findings could be due to the wide range of terms used to describe self-esteem and self-concept (Liu et al., 2015). As recommended by Biddle and Asare (2011), we included self-esteem as well as its physical domains, including physical self-concept, body image, and physical self-worth, as these may be more proximal targets for PA interventions (Babic et al., 2014; Biddle & Asare, 2011). Two reviews found inconsistencies in how self-esteem, physical self-concept, and its sub-domains were defined and assessed (Babic et al., 2014; Liu et al., 2015). Again, more research is needed to determine how the “self” variables are affected by PA, in particular the relationship between self-esteem and physical self-concept. As physical self-concept is considered to be a mediator between PA and self-esteem (Liu et al., 2015), improvements in physical self-concept should positively impact self-esteem when those improvements are valued (Babic et al., 2014; Lubans et al., 2016); however, this was not observed in all reviews (Liu et al., 2015). This is not unexpected theoretically given the many other sources that could contribute to self-esteem (e.g. academic performance; social relationships). Self-esteem is a multidimensional mental health outcome and additional mediator and moderator

analyses are needed to determine what intervention components work best, and for whom, to improve physical self-concept and self-esteem (Babic et al., 2014; Liu et al., 2015).

3.4. Strengths and limitations

Our research question and inclusion criteria allowed for a broad scope, which is one reason we chose to conduct an umbrella review (Aromataris et al., 2015). This format also allowed us to compare and contrast findings from recent reviews published from 2010 onwards and identify areas of future research. The use of AMSTAR 2 to assess the quality of included reviews was a strength of our umbrella review. Another strength of our umbrella review was the inclusion of reviews on children, as many mental health reviews have included adolescents only, as well as including longitudinal studies to determine if PA in childhood could have potential protective effects on mental health in adolescence. With the exception of one review (Poitras et al., 2016), each review that examined longitudinal studies found associations between PA and lower depression (Dennison et al., 2016; Hoare et al., 2014; Korczak et al., 2017), lower anxiety (Eime et al., 2013), higher self-esteem (Eime et al., 2013) and higher physical self-concept (Babic et al., 2014); however, the generally short follow-up (1–2 years) indicate more longitudinal studies are needed.

We acknowledge the limitations to our review. We did not submit our umbrella review protocol for peer-review; however, the review protocol was reviewed, modified, and agreed upon by the *Expert Statement* panel. We made peer-reviewed publication restrictions to ensure rigor of the evidence; however, as all of our included reviews also made this restriction, the findings may be susceptible to publication and selective reporting biases. Restrictions to the English language were made for practical reasons (e.g. insufficient budget for article translation), which also means some reviews on the topic may have been missed. A potential limitation of umbrella reviews is overlap in studies, although our assessment of the number of new unique studies (n = 150) indicated that this was not a large concern. A major limitation to the findings in our umbrella review was the overall low quality of the reviews, as well as the low quality of many of the individual studies included in the reviews.

4. Conclusions

It is imperative to identify and implement effective interventions to reduce the public health burden of mental illness in children and youth. PA has been shown to have positive mental health outcomes in children and youth. Specifically, the strongest evidence appears to be depression/depressive symptoms, in particular for those already diagnosed with depression, although all participants would likely benefit from PA. The inclusion of several meta-analyses examining intervention effects have significantly advanced this field. There was also evidence supporting the effectiveness of PA to improve the physical self-concept domain of self-esteem. Notably in terms of testing potential causal mechanisms in a recent systematic review, the strongest evidence has been found for improvements in physical self-perceptions that accompanied enhanced self-esteem in the majority of studies measuring these outcomes (Lubans et al., 2016). There is less certainty about the role of PA in reducing anxiety among children and youth and research focus in this area appears to have stagnated. Overall, PA appears to be an effective and low-risk intervention that can augment current therapies, although more high-quality research, using robust designs, as well as additional moderator analyses are needed to determine what type of PA intervention may result in better mental health outcomes for children and youth.

Acknowledgements

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. Guy Faulkner is supported by a Canadian Institutes of Health Research-Public Health Agency of Canada Chair in Applied Public Health. Leigh Vanderloo is supported by a Canadian Institutes of Health Research Fellowship Award.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.mhpa.2018.12.001.

Conflicts of interest

The authors have no conflicts to declare.

References

* = review included in the umbrella review.

Ahn, S., & Fedewa, A. L. (2011). A meta-analysis of the relationship between children's physical activity and mental health. *Journal of Pediatric Psychology, 36*(4), 385–397.

Aromataris, E., Fernandez, R., Godfrey, C. M., Holly, C., Khalil, H., & Tungpunkom, P. (2015). Summarizing systematic reviews: Methodological development, conduct and reporting of an umbrella review approach. *International Journal of Evidence-based Healthcare, 13*(3), 132–140. https://doi.org/10.1097/XEB.0000000000000055.

*Ash, T., Bowling, A., Davison, K., & Garcia, J. (2017). Physical activity interventions for children with social, emotional, and behavioral disabilities—a systematic review. *Journal of Developmental and Behavioral Pediatrics, 38*(6), 431–445.

*Babic, M. J., Morgan, P. J., Plotnikoff, R. C., Lonsdale, C., White, R. L., & Lubans, D. R. (2014). Physical activity and physical self-concept in youth: Systematic review and meta-analysis. *Sports Medicine, 44*(11), 1589–1601. https://doi.org/10.1007/s40279-014-0229-z.

*Bailey, A. P., Hetrick, S. E., Rosenbaum, S., Purcell, R., & Parker, A. G. (2017). Treating depression with physical activity in adolescents and young adults: A systematic review and meta-analysis of randomised controlled trials. *Psychological Medicine, 47*(8), 1068–1083.

*Bassett-Gunter, R., McEwan, D., & Kamarhie, A. (2017). Physical activity and body image among men and boys: A meta-analysis. *Body Image, 22*, 114–128. https://doi.org/10.1016/j.bodyim.2017.06.007.

Biddle, S. J. H., & Asare, M. (2011). Physical activity and mental health in children and adolescents: A review of reviews. *British Journal of Sports Medicine, 45*(11), 886–895.

Boak, A., Hamilton, H. A., Adlaf, E. M., Henderson, J. L., & Mann, R. E. (2018). *The mental health and well-being of Ontario students, 1991-2017*. Toronto, ON: Centre for Addictions and Mental Health.

*Brown, H. E., Pearson, N., Braithwaite, R. E., Brown, W. J., & Biddle, S. J. (2013). Physical activity interventions and depression in children and adolescents: A systematic review and meta-analysis. *Sports Medicine, 43*(3), 195–206. https://doi.org/10.1007/s40279-012-0015-8.

*Burkhardt, J., & Brennan, C. (2012). The effects of recreational dance interventions on the health and well-being of children and young people: A systematic review. *Arts & Health: International Journal of Research in Pharmacology & Pharmacotherapeutics, 4*(2), 148–161. https://doi.org/10.1080/17533015.2012.665810.

*Burnsall, P. (2014). The relationship between physical activity and depressive symptoms in adolescents: A systematic review. *Worldviews on Evidence-based Nursing, 11*(6), 376–382.

*Carter, T., Morres, I. D., Meade, O., & Callaghan, P. (2016). The effect of exercise on depressive symptoms in adolescents: A systematic review and meta-analysis. *Journal of the American Academy of Child & Adolescent Psychiatry, 55*(7), 580–590. https://doi.org/10.1016/j.jaac.2016.04.016.

Catalan-Matamoros, D., Gomez-Conesa, A., Stubbs, B., & Vancampfort, D. (2016). Exercise improves depressive symptoms in older adults: An umbrella review of systematic reviews and meta-analyses. *Psychiatry Research, 244*, 202–209. https://doi.org/10.1016/j.psychres.2016.07.028.

*Dennison, M., Sisson, S. B., & Morris, A. (2016). Obesogenic behaviours and depressive symptoms in children: A narrative literature review. *Obesity Reviews, 17*(8), 735–757.

*Eime, R. M., Young, J. A., Harvey, J. T., Charity, M. J., & Payne, W. R. (2013). A systematic review of the psychological and social benefits of participation in sport for children and adolescents: Informing development of a conceptual model of health through sport. *International Journal of Behavioral Nutrition and Physical Activity, 10*, 98.

Ekeland, E., Heian, F., & Hagen, K. B. (2005). Can exercise improve self esteem in children and young people? A systematic review of randomised controlled trials. *British Journal of Sports Medicine, 39*(11), 792. https://doi.org/10.1136/bjsm.2004.017707.

*Evans, M. B., Allan, V., Erickson, K., Martin, L. J., Budziszewski, R., & Cote, J. (2017). Are all sport activities equal? A systematic review of how youth psychosocial experiences vary across differing sport activities. *British Journal of Sports Medicine, 51*(3), 169–176. https://doi.org/10.1136/bjsports-2016-096725.

Faulkner, G., & Duncan, M. (2018). Physical activity and mental health: A focus on depression. In J. Piggin, L. Mansfield, & M. Weed (Eds.). *Routledge handbook of physical activity policy and practice*. Abingdon, OX: Routledge.

Gunnell, K. E., Poitras, V. J., LeBlanc, A., Schibli, K., Barbeau, K., Hedayati, N., et al.

(2018). *Physical activity and brain structure, brain function, and cognition in children and youth: A systematic review of randomized controlled trials. Mental health and physical activity*. https://doi.org/10.1016/j.mhpa.2018.11.002.

∗Hermens, N., Super, S., Verkooijen, K. T., & Koelen, M. A. (2017). A systematic review of life skill development through sports programs serving socially vulnerable youth. *Research Quarterly for Exercise & Sport, 88*(4), 408–424.

∗Hoare, E., Skouteris, H., Fuller-Tyszkiewicz, M., Millar, L., & Allender, S. (2014). Associations between obesogenic risk factors and depression among adolescents: A systematic review. *Obesity Reviews, 15*(1), 40–51.

∗ten Hoor, G. A., Kok, G., Peters, G. Y., Frissen, T., Schols, A. M. W. J., & Plasqui, G. (2017). The psychological effects of strength exercises in people who are overweight or obese: A systematic review. *Sports Medicine, 47*(10), 2069–2081. https://doi.org/10.1007/s40279-017-0748-5.

∗Janssen, I., & Leblanc, A. G. (2010). Systematic review of the health benefits of physical activity and fitness in school-aged children and youth. *International Journal of Behavioral Nutrition and Physical Activity, 7*, 40.

∗Johnson, K. E., & Taliaferro, L. A. (2011). Relationships between physical activity and depressive symptoms among middle and older adolescents: A review of the research literature. *Journal for Specialists in Pediatric Nursing: Journal for Specialists in Pediatric Nursing, 16*(4), 235–251.

∗Joronen, K., Aikasalo, A., & Suvitie, A. (2017). Nonphysical effects of exergames on child and adolescent well-being: A comprehensive systematic review. *Scandinavian Journal of Caring Sciences, 31*(3), 449–461.

Kandola, A., Vancampfort, D., Herring, M., Rebar, A., Hallgren, M., Firth, J., et al. (2018). Moving to beat anxiety: Epidemiology and therapeutic issues with physical activity for anxiety. *Current Psychiatry Reports, 20*(8), 63.

∗Korczak, D. J., Madigan, S., & Colasanto, M. (2017). Children's physical activity and depression: A meta-analysis. *Pediatrics, 139*(4), 1–14.

Liberati, A., Altman, D., Tetzlaff, J., Mulrow, C., Gotzsche, P., Ioannidis, J., ... Moher, D. (2009). The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: Explanation and elaboration. *PLoS Medicine, 6*(7), e1000100.

∗Liu, M., Wu, L., & Ming, Q. (2015). How does physical activity intervention improve self-esteem and self-concept in children and adolescents? Evidence from a meta-analysis. *PLoS One, 10*(8), e0134804.

∗Lubans, D. R., Plotnikoff, R. C., & Lubans, N. J. (2012). A systematic review of the impact of physical activity programmes on social and emotional well-being in at-risk youth. *Child and Adolescent Mental Health, 17*(1), 2–13. https://doi.org/10.1111/j.1475-3588.2011.00623.x.

∗Lubans, D., Richards, J., Hillman, C., Faulkner, G., Beauchamp, M., Nilsson, M., ... Biddle, S. (2016). Physical activity for cognitive and mental health in youth: A systematic review of mechanisms. *Pediatrics, 138*(3).

∗Poitras, V. J., Gray, C. E., Borghese, M. M., Carson, V., Chaput, J., Janssen, I., ... Tremblay, M. S. (2016). Systematic review of the relationships between objectively measured physical activity and health indicators in school-aged children and youth. *Applied Physiology Nutrition and Metabolism, 41*(6 Suppl 3), S197–S239.

Polanczyk, G. V., Salum, G. A., Sugaya, L. S., Caye, A., & Rohde, L. A. (2015). Annual research review: A meta-analysis of the worldwide prevalence of mental disorders in children and adolescents. *Journal of Child Psychology and Psychiatry, 56*(3), 345–365. https://doi.org/10.1111/jcpp.12381.

∗Radovic, S., Gordon, M. S., & Melvin, G. A. (2017). Should we recommend exercise to adolescents with depressive symptoms? A meta-analysis. *Journal of Paediatrics and Child Health, 53*(3), 214–220.

∗Rafferty, R., Breslin, G., Brennan, D., & Hassan, D. (2016). A systematic review of school-based physical activity interventions on children's wellbeing. *International Review of Sport and Exercise Psychology, 9*(1), 215–230. https://doi.org/10.1080/1750984X.2016.1164228.

Rebar, A. L., Stanton, R., Geard, D., Short, C., Duncan, M. J., & Vandelanotte, C. (2015). A meta-meta-analysis of the effect of physical activity on depression and anxiety in non-clinical adult populations. *Health Psychology Review, 9*(3), 366–378. https://doi.org/10.1080/17437199.2015.1022901.

Rosenbaum, S., Tiedemann, A., Sherrington, C., Curtis, J., & Ward, P. B. (2014). Physical activity interventions for people with mental illness: A systematic review and meta-analysis. *Journal of Science and Medicine in Sport, 18*. https://doi.org/10.1016/j.jsams.2014.11.161 e150-e150.

∗Ruotsalainen, H., Kyngas, H., Tammelin, T., & Kaariainen, M. (2015). Systematic review of physical activity and exercise interventions on body mass indices, subsequent physical activity and psychological symptoms in overweight and obese adolescents. *Journal of Advanced Nursing, 71*(11), 2461–2477. https://doi.org/10.1111/jan.12696 [doi].

Sallis, J., Prochaska, J., & Taylor, W. (2000). A review of correlates of physical activity of children and adolescents. *Medicine & Science in Sports & Exercise, 32*(5), 963–975. https://doi.org/10.1097/00005768-200005000-00014.

Schuch, F. B., Deslandes, A. C., Stubbs, B., Gosmann, N. P., da Silva, C. T. B., & Fleck, M. P. d. A. (2017). Factors that influence the neurobiological effects of exercise likely extend beyond age and intensity in people with major depression. *Neuroscience & Biobehavioral Reviews, 77*, 301–302.

Shea, B. J., Reeves, B. C., Wells, G., Thuku, M., Hamel, C., Moran, J., ... Henry, D. A. (2017). AMSTAR 2: A critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. *BMJ, 358*, j4008. https://doi.org/10.1136/bmj.j4008.

∗Spruit, A., Assink, M., van Vugt, E., van der Put, C., & Stams, G. J. (2016). The effects of physical activity interventions on psychosocial outcomes in adolescents: A meta-analytic review. *Clinical Psychology Review, 45*, 56–71.

Tremblay, M. S., Carson, V., Chaput, J., Connor Gorber, S., Dinh, T., Duggan, M., ... Poitras, V. J. (2016). Canadian 24-hour movement guidelines for children and youth: An integration of physical activity, sedentary behaviour, and sleep. *Applied Physiology Nutrition and Metabolism, 41*, S311–S327.

Vanderloo, L. M., Cairney, J., Choquette, L., Collet, J. P., Davies, T., Faulkner, G., et al. (2018). *Expert statement on physical activity and brain health in children and youth.* https://www.participaction.com/sites/default/files/downloads/expert_statement_on_physical_activity_and_brain_health_in_children_and_youth.pdf/, Accessed date: 30 August 2018.

Warburton, D. E. R., Nicol, C. W., & Bredin, S. S. D. (2006). Health benefits of physical activity: The evidence. *Canadian Medical Association Journal, 174*(6), 801–809.