McMichan, L., Gibson, A-M., & Rowe, D. A. (2017). Classroom-based physical activity and sedentary behavior interventions in adolescents: a systematic review and meta-analysis. *Journal of Physical Activity and Health*.

Abstract

Background: It is reported that 81% of adolescents are insufficiently active. Schools play a pivotal role in promoting physical activity (PA) and reducing sedentary behavior (SB). The aim of this systematic review and meta-analysis was to evaluate classroom-based PA and SB interventions in adolescents. **Methods:** A search strategy was developed using the PICOS framework. Articles were screened using strict inclusion criteria. Study quality was assessed using the EPHPP quality assessment tool (http://www.ephpp.ca/tools.html). Outcome data for pre- and post- intervention were extracted and effect sizes were calculated using Cohen's d. **Results:** The strategy yielded 7574 potentially relevant articles. Nine studies were included for review. Study quality was rated as strong for one study, moderate for five studies and weak for three studies. Five studies were included for meta-analyses, which suggested that the classroom-based interventions had a non-significant effect on PA ($p = 0.55$, $d = 0.05$) and a small, non-significant effect on SB ($p = 0.16$, $d = -0.11$). **Conclusion:** Only nine relevant studies were found and the effectiveness of the classroom-based PA and SB interventions varied. Based on limited empirical studies, there is not enough evidence to determine the most effective classroom-based methodology to increase PA and SB.

Introduction

Physical activity (PA) plays an important role in adolescents' health. Being active has been shown to benefit physiological[1] and psychological[1,2] health. Recent evidence shows that physical inactivity and sedentary behavior (SB) are highly prevalent amongst adolescents.[3] Schools can play a role in improving PA, with government organizations highlighting their importance and adding policies into their individual frameworks.[4,5] Research into school-based PA and SB interventions has increased in recent years and a number of reviews have established the efficacy of such interventions.[6-13] School-based interventions have been shown to increase moderate-to-vigorous physical activity (MVPA)[8,9] and VO_{2max} (a measure of cardiovascular fitness) and have also had a positive effect on television viewing (a proxy measure of SB)[6], yet some reviews have found inconclusive evidence that such interventions have an effect on overall PA.[8,12] Few reviews have looked at the effects of school-based interventions on SB. Hynynen et al. (2015) analyzed four studies that measured SB and reported that only two had shown significant decreases in SB. This indicates that there is little research into school-based interventions that target reducing SB.

Interventions to increase PA and reduce SB within the classroom are fairly novel. For the purpose of this review, interventions were delimited to those conducted in traditional classrooms that were not physical activity specific (i.e. physical education interventions were excluded). Classroom-based PA and SB interventions have a number of potential benefits other than improving PA and reducing SB[9], such as improving on-task behavior[14] and academic performance.[15,16] However, most classroom-based research has focused on primary/elementary school children rather than secondary/middle/ high school adolescents.

In the development of interventions, evaluating the effectiveness of the intervention on the desired outcomes is important.[13] There are different evaluation frameworks, such as the RE-AIM framework.[17] Implementation is one of the RE-AIM factors that determines

whether the intervention was delivered as intended.[17] According to Durlak and Dupre[18], implementation includes fidelity, dosage, quality, participant responsiveness, program differentiation, monitoring of controls, program reach and adaptation.[18] All of these aspects of implementation are important in establishing the validity of interventions, however the reporting of implementation appears to be rare, particularly for school-based PA and SB interventions.[13] Naylor et al.[19] systematically reviewed implementation in school-based PA interventions. Of the 15 studies included, 11 suggested positive associations between health outcomes and level of implementation. Implementation elements (e.g., fidelity, dosage) were measured using various measurement tools and/or techniques. The literature on the role of implementation and the intervention effectiveness appears scarce. Further research should assess implementation in relation to outcomes.[19]

Several systematic reviews have examined school-based interventions, of which only one has focused on classroom-based PA interventions.[9] Only one of the original studies included in that review was based in a secondary school[20], however PA was not an outcome measure. Therefore, the aims of this current systematic review were to: 1) review classroom-based PA and SB interventions within an early secondary/ middle/ high school setting and determine the most effective methodology for increasing PA and reducing SB; and 2) determine if implementation has an impact on the effectiveness of the interventions.

Research has indicated that there are psychological constructs which are correlates towards PA in adolescents (e.g. self-efficacy, autonomy)[21], therefore a secondary aim is to determine if these interventions change any psychological constructs, and if these changes effect PA behavior.

Methods

This systematic review protocol was registered and published under Prospero [CRD42015026721] in October 2015. The protocol was constructed using the guidelines in

the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement.[22,23]

Search Strategy

The following six electronic databases were searched: Medline (OVID); EMBASE; ERIC; SportDiscus; PsycInfo; and Web of Science. The same search strategy was used for each database, with adaptations of wildcards/truncation symbols to fit the criteria for each specific database. The search strategy was conducted in October 2015 and was cross-checked in November 2015. The cross-checking involved the second and third authors, who conducted the search strategy for each database at the same time to ensure consistency. A search of the grey literature was performed via the Open Grey database (http://www.opengrey.eu/). In July 2017, the search strategy was conducted again. This strategy was adapted to broaden the number of articles retrieved to ensure no articles were missed. The adaptations included the addition of 'child*' and 'lesson*' to further enhance the search.

The search strategy was determined using the PICOS framework, and is presented in Table 1. Physical education was not included in the search strategy as this review is focused on classroom-based PA and SB programs beyond the physical education setting.

Study Selection

Following the search strategy and retrieval of references, these were exported into EndNote Reference Manager, version X6 (Thomson Reuters, Philadelphia). Duplicates were also removed via the EndNote Reference Manager software. Duplicates were visually inspected to ensure the correct references were removed. References included for screening were exported into a Microsoft Excel, version 2013 (Micrososoft Corp, Redmond, WA) spreadsheet. Manual inspection of duplicates was performed again to ensure that there were no duplicates.

104 The inclusion criteria for screening articles were: a) randomized controlled trials
105 (RCTs), controlled trials (CTs), quasi-experimental, or pre- and post- study designs; b)
106 studies based in a classroom setting only, which targeted PA or SB, or both PA and SB; c)
107 non-clinical secondary/ middle/ high school adolescents between the ages of 11-15 years old.
108 We excluded any study based in primary/elementary schools or in high/secondary schools
109 where interventions targeted adolescents over the age of 15 years. This is because primary
110 and secondary school environments are different in terms of education and the structure of
111 the school day.

112 The exclusion process involved reviewing titles of the articles that were generated
113 from the search strategy. Titles that did not match the criteria (e.g. clinical populations,
114 outside school hours) were excluded. Article titles, which were potentially relevant, were
115 then reviewed at abstract level. Abstracts of articles, which appeared to meet the inclusion
116 criteria, were then reviewed at full-text. For abstracts and full-texts where there was
117 uncertainty, the second author (AMG) cross-checked for confirmation. Any discrepancies
118 were subsequently discussed in consultation with the third author (DAR) until a decision was
119 agreed. Reference lists from review and summary articles that were retrieved from the search
120 were checked to ensure that no articles were missed.

121 **Data Collection Process and Data Items**

122 The following data were extracted and entered into a standardized form in Microsoft
123 Word, version 2013 (Micrososoft Corp, Redmond, WA): author(s); date of publication;
124 country the study was conducted in; aim of the study; study design; population; intervention;
125 and results of the intervention.

126 **Risk of Bias in Individual Studies**

127 An adapted version of the Effective Public Health Practice Project (EPHPP)[24] tool
128 was used for quality assessment. The EPHPP has a rating scale of 1 to 3 (1 = strong, 2 =

moderate, 3 = weak) and the quality was assessed on selection bias, study design, confounders, blinding, data collection methods, and withdrawal and drop-outs.[24] Selection bias was scored based on population representativeness, and percentage agreeing to take part. Study design was scored on the type of design used. Strong was awarded if the studies were a randomized control trial or control clinical trial. The authors adapted this to include group/cluster randomized control trials, as previously adapted by Chillon et al.[25] This adaptation was made due to the nature of school-based interventions whereby schools and/or classes are often randomized rather than the individuals. Confounders was scored on differences between groups at baseline, and the percentage of confounders controlled. Blinding was scored based on whether the participants were blinded to the research question, and the assessors were blinded to the group allocation. The authors added in a 'not relevant' option to this category. This decision was made because blinding might not be possible within a school setting, especially if classes are randomized. Pupils are unlikely to be aware of the research question itself; however, they may have an understanding of why the study is taking place. Data collection was scored based on the evidence reported for validity and reliability of the measurement tools used. Finally, withdrawal and dropout was scored on the percentage of participants completing the study. A global rating was then determined based on the ratings of the above constructs. A strong global rating was awarded if no weak ratings were present, moderate global rating if there was only one weak rating and a weak global rating if there were two or more weak ratings.[24] Intervention integrity (assessed for whether the intervention consistency was measured; what percentage received the intervention; was there potential for contamination) and appropriate analysis in relation to the research question(s) (unit of analysis; unit of allocation; statistical analysis; intention to treat) were also assessed. However, the scoring of these constructs did not contribute to the overall rating score.

154 **Summary of Measures**

155 The primary outcome measures were PA and SB. Where possible, pre- and post- data
156 were extracted from both the intervention and control groups, and was inputted into
157 Microsoft Excel, version 2013 (Micrososoft Corp, Redmond, WA). Means and standard
158 deviations (SD) were extracted from each study. If SDs were not reported directly, they were
159 calculated based on reported standard errors and sample sizes.[26] Cohen's d effect sizes were
160 calculated from means and SDs to determine the interaction effect, and where an interaction
161 effect could not be determined (i.e., if the study had no control condition or if the study only
162 reported post-intervention data), the effect size was calculated using pre- and post-
163 intervention data only, or post-intervention data only. The effect sizes were interpreted as
164 small (d = 0.2), medium (d = 0.5), or large (d = 0.8), following the guidelines of Cohen.[27]

165 **Meta-Analysis**

166 A meta-analysis was performed to determine the overall effect of classroom-based
167 interventions on PA and SB. Review Manager, version 5.3 (RevMan 5.3) (The Nordic
168 Cochrane Centre, Copenhagen)[28] computer software was used to conduct the meta-analysis.
169 Only studies that used a two group (intervention/control), pre-post design were included in
170 the meta-analysis (n = 5). Data from baseline measures and the first measurement post-
171 intervention were analyzed. The data inputted into RevMan 5.3 were: the standardized mean
172 differences between pre- and post-intervention for the intervention group, and the control
173 group; the pooled SD of the four cells of data (pre- and post- intervention data, pre- and post-
174 control data); and the sample size of each group (*n*). This produced a Cohen's d for the
175 interaction effect and 95% confidence interval (CI). The RevMan 5.3 software then pooled
176 the effects for all studies to produce an overall effect, weighted by individual study sample
177 size. Standardized means were calculated to take into account that each study used different
178 measures of PA/SB. Due to the heterogeneity of the studies (I^2), a random-effects model was

used for the analysis, and standardized mean differences were used to account for the different measurement outputs from the studies.

Results

Study Selection

Initial search strategies yielded 7574 potentially relevant articles. 1767 duplicates were removed. 5556 studies were excluded during the title and abstract screening stages, 242 were excluded at full text level, leaving nine studies included for the systematic review, [29-37] and five included for the meta-analysis.[31,33,34,35,37] A summary of the screening process along with reasons for full text exclusions is shown in Figure 1.

Study Characteristics

Five studies were based in the USA,[29,30,32,34,37] two were based in China[31,33], one was based in the UK[36] and one was based in Iran.[33] Seven studies were cluster randomized control trials[30,31,33-37] and two were pre- and post- cohort design with no control group.[29,32] Sample size ranged from N = 85[33] to N = 1391.[36] The reported mean age of participants ranged from 12.0 years to 15.3 years. Whittermore et al.[37] included adolescents who were 16-17 years (~30% of the sample). This study was still included on the basis that ~70% of the sample met our age range criterion and the study was conducted within a secondary/high school setting. One study did not report mean age but stated the intervention was targeted to years 7 to 9, which would correspond to an age range of 11-14 years in the English secondary school system.[36]

Interventions

The interventions included in the study were all classroom-based and were educational. Three of the studies investigated PA as an outcome[33,34,36] and one study investigated both PA and SB as outcomes[35]. Five studies had a nutritional element to the

203 program alongside PA and SB.[29-32,37] No studies investigated reducing SB only. The
204 nutritional elements included, for example, education on fruit and vegetable consumption[29]
205 and measuring these outcome variables. Five studies measured psychological outcomes
206 including self-efficacy, motivation, and attitudes. [29,30,33,35,37] Details on the interventions are
207 presented in Table 2.

208 **Theoretical Underpinnings**

209 Six of the eight studies reported using one or more theoretical frameworks to inform
210 their interventions. Two used Self-Determination Theory[30,35,38]; two used Social Cognitive
211 Theory [30,31,39]; one used the Theory of Planned Behavior[29,40]; one used Social Learning
212 Theory[37,41]; one used Theory of Meanings of Behaviour,[35,42] and one study used a version of
213 Stages of Change model.[43]

214 **Physical Activity/ Sedentary Behavior**

215 All PA data were collected through self-reported measures, except for one study in
216 which PA was measured objectively[36]. Varieties of PA outcome measures were reported.
217 These included: MVPA (mins/day)[31,36]; PA performance[33]; PA expressed as the number of
218 30-minute blocks spent in each of three intensities (high, medium, light)[35]; PA (days/week)[32];
219 moderate exercise (days/week for at least 30 minutes); vigorous exercise (day/week for 20
220 minutes)[37]; PA frequency (presented as a score of 1-4; 1 = never, 2 = 2 times per week, 3 = 3-
221 4 times per week, 4 = almost every day)[34]; walking; and stair climbing.[29,30] There were varied
222 results regarding the effects of the interventions on PA behavior, with only three studies
223 reporting significant results. Contento et al.[30] found a significant increase in walking for
224 transport and walking for exercise (0.55 days/week, $p < 0.001$, $d = 0.26$; 0.36 days/week, $p =$
225 0.044, $d = 0.14$, respectively) compared to control post intervention. Dunton et al.[32] reported
226 an increase of 0.43 days/week ($p < 0.001$, $d = 0.2$) at post intervention compared to baseline.

There were no significant differences in PA frequency score reported by Schwarzer et al.[34] when all participants were analyzed together. However, when participants were split into Stages of Change (preintenders = low intention of performing PA; intenders = those who intend on performing PA; actors = those who perform PA), the highest increase in PA frequency score was found in preintenders (those least likely to take part in PA). This group had a significant increase PA frequency score of 0.84 ($p < 0.01$, $d = 1.23$), raising their score from 2.08 ± 0.60 at baseline to 2.92 ± 0.76 post intervention in the resource communication group. This was higher than in the planning intervention group (2.15 ± 0.71 vs 2.60 ± 0.92, respectively) which was non-significant ($p > 0.05$, $d = 0.55$).

Six studies[33,34] measured outcomes of SB, using self-report. One study measured SB (mins/day)[31]; one measured SB (hours/day)[37]; one measured screen time (television viewing/game play/internet usage) in hours/day[32]; one measured screen time in half hour blocks[35]; two studies measured screen time in days/week[29,30]. Four studies reported significant decreases in SB. Dunton et al.[32] reported a significant decrease in time playing video games/computer use (0.31 hours/day; $p = 0.002$, $d = -0.21$) and time watching television (0.16 hours/day; $p = 0.024$, $d = -0.15$) post intervention. Contento et al.[29] reported significant decreases in the number of days pupils watched television and played video games (0.33 days/week, $p = 0.003$, $d = -0.18$; 0.60 days/week, $p < 0.001$, $d = -0.25$, respectively). Contento et al.[30] reported a significant ($p < 0.001$, $d = -0.38$) decrease in leisure screen time (days/week) in the intervention group compared to control post intervention (4.85 ± 1.8 vs 5.51 ± 1.7 days/week, respectively). Spruijt-Metz et al.[35] reported a significant decrease in screen time in the intervention group compared to the control ($p < 0.05$, $d = -0.28$).

Whittemore et al.[37] reported significant differences between baseline, 3 month and 6 month follow up in vigorous PA (hours/day) in both the HEALTH[e]TEEN (control) and HEALTH[e]TEEN + Coping Skills Training (CST) groups ($p < 0.01$, $d = 0.032$; $p < 0.01$, $d =$

252 0.031, respectively) and, SB weekday (hours/day) ($p < 0.01$, $d = -0.25$; $p < 0.01$, $d = -0.31$)
253 and SB weekends ($p < 0.01$, $d = -0.35$; $p < 0.01$, $d = -0.31$). Only the HEALTH[e]TEEN +
254 CST had a significant increase on moderate PA (HEALTH[e]TEEN + CST $p < 0.01$, $d =$
255 0.27; HEALTH[e]TEEN $p = 0.06$, $d = 0.18$). However, the difference between the two
256 groups were non-significant for moderate and vigorous PA.

257 **Psychological Outcomes**

258 Psychological outcomes were measured in five of the studies.[29,30,33,35,37] Three studies
259 measured pupils' self-efficacy.[29,30,37] Two studies measured self-efficacy of walking and stair
260 climbing.[29,30] Contento et al.[29] reported a significant ($p = 0.008$, $d = 0.2$) increase in self-
261 efficacy for stair climbing from baseline to post intervention (3.70 ± 1.78 vs 4.00 ± 1.08,
262 respectively) although there was no significant change in self-efficacy for walking ($p = 0.42$,
263 $d = 0.08$). Contento et al.[30] reported a significant difference between intervention and control
264 post-intervention for self-efficacy for walking and stair climbing (combined) (2.89 ± 0.77 vs
265 2.60 ± 0.81, $p < 0.001$, $d = 0.37$, respectively). Whittemore et al.[37] reported that both groups
266 (HEALTH[e]TEEN vs HEALTH[e]TEEN + CST) significantly ($p < 0.01$, $d = 0.26$; $p < 0.01$,
267 $d = 0.33$) increased self-efficacy for exercise from baseline to follow up, yet there were no
268 significant differences between the two groups ($p = 0.46$, $d = 0.08$).

269 Motivation was measured in two studies.[30,35] Contento et al.[30] measured pupils'
270 autonomous motivation and reported significantly ($p = 0.005$) higher autonomy and
271 competence towards PA in the intervention groups compared to the control group (autonomy
272 = 3.13 ± 0.74 vs 2.94 ± 0.82, $d = 0.24$; competence = 3.13 ± 0.77 vs 2.95 ± 0.88, respectively,
273 $d = 0.22$). Spruijt-Metz et al.[35] measured the different constructs of motivation (external
274 regulation, introjected regulation, identified regulation, and intrinsic motivation), with the
275 exception of amotivation. Intrinsic motivation was the only form of motivation that

significantly changed. The control group started off with higher scores compared to the intervention at baseline (1.24 ± 0.06 vs 1.11 ± 0.07, respectively) yet post intervention, there was a decrease in the control and an increase in intervention (1.18 ± 0.06 vs 1.16 ± 0.07, respectively), corresponding to a net effect of 0.11 (d = 0.11) in favor of the intervention group.

Two studies measured attitudes towards physical activity[33] and walking.[29] Ghaffari[33] reported a significant ($p < 0.001$) increase in attitude scores from baseline to post intervention (d = 1.71) and follow up (46.47 ± 3.43 vs 53.94 ± 2.11; d = 1.71 vs 52.07 ± 4.06; d = 0.88, respectively). The post intervention and one-month follow up scores in the intervention were significantly higher than in the control group (53.94 ± 2.11 vs 47.58 ± 5.76 for post intervention scores respectively; 52.07 ± 4.06 vs 49.72 ± 4.27 for one month follow up, respectively). Attitudes towards walking significantly increased after the intervention by Contento et al.[29] compared to baseline scores (4.16 ± 0.73 vs 4.30 ± 0.69, respectively, $p = 0.022$, d = 0.2). A full summary is presented in Table 2.

Quality Assessment

Quality assessment was performed on the nine studies included. Of the nine studies, one was rated as strong[35], five were rated as moderate[29-31,36,37] and three were rated as weak[32-34]. A summary of the ratings for each category is presented in Table 3.

Implementation

Five studies reported monitoring of implementation.[29-31,36,37] To ensure fidelity, Contento et al.[29] had a member of the research team observe at least one class per week, provided all materials, and met weekly with teaching staff to provide guidance on how the lessons should be run. Similarly, Contento et al.[30] had two members of the research team attend one third of lessons taught by each teacher, provided guidance on how the lessons

300 should be run and provided all materials. Whittemore et al.[37] consulted with teachers prior to
301 the intervention to try to optimize implementation. Pupil participation was also monitored by
302 the research team bi-monthly.

303 Cui et al.[31] monitored implementation through direct observation. A research member
304 and an external figure observed the peer education classes. Immediately post intervention, a
305 focus group was conducted with pupils and interviews were conducted with staff members of
306 the schools. Results of the observation suggested that the material and classes delivered by
307 the peer leaders met the content and objectives that were presented in the peer leaders'
308 manual. The data collected through interviews with staff members indicated that the
309 intervention was feasible and acceptable.

310 Tymms et al.[36] monitored implementation also by direct observation. Researchers
311 were present for one or more classes and these classes were scored on how much they
312 adhered to the program. The researchers also followed up with questionnaires (teachers and
313 students) and focus groups (students).

314 **Meta-Analysis**

315 The results of the random-effects meta-analysis showed there were no significant
316 individual study effects on PA or SB for the interventions included in the analysis. For PA,
317 the overall effect of the interventions across the five included studies was non-significant ($p =$
318 0.55, $d = 0.05$, 95% CI [-0.11, 0.21]). For SB, the overall effect of the interventions across
319 the three included studies was non-significant ($p = 0.16$, $d = -0.11$, 95% CI [-0.25, 0.04]). The
320 I^2 for both meta-analyses indicated that there was substantial heterogeneity of the studies
321 (67% and 52% for PA and SB respectively). The I^2 percentage determines the variance that
322 could be attributed to the heterogeneity of the studies included for analysis. Forest plots of
323 the meta-analyses are presented in Figure 2 and 3.

Discussion

The aims of this systematic review were to: 1) review classroom-based PA and SB interventions within an early secondary/ middle/ high school setting and determine the most effective methodology for increasing PA and reducing SB; 2) determine if implementation has an impact on the effectiveness of the interventions; and 3) determine if these interventions have an impact on psychological constructs. Nine studies were included for review based on the inclusion criteria. These studies varied considerably in design, and the interventions had varying effects on PA, SB and psychological outcomes.

Summary of Evidence

Overall, the evidence collated from the review and meta-analysis has shown that classroom-based PA and SB interventions in early secondary schools have yielded mostly small or no effects on PA and SB. Results of both meta-analyses were non-significant. For studies that found significant effects on PA[29, 32], these effects were only small, and were based on post-intervention data only[29] and single group, pre- and post- intervention data[32]. The study by Schwarzer et al.[34] did find a significant large effect for PA in the resource communication group for preintenders (d = 0.96) however, when all stages were analyzed together, this effect was non-significant and only a small effect (d = 0.3). This stage of change is similar to the stage of 'precontemplation' in the more commonly known Transtheoretical Model (TTM).[44] The authors used their own Stages of Change classifications (preintenders, intenders, and actors) instead of more traditional models such as the TTM. The resource communication intervention focused on the advantages and disadvantages of being physically active and being sedentary.

Four studies reported significant decreases in outcome measures for SB[29,30,31,35] but these were only small effects. Importantly, two of these studies lacked a control group[29,32]

348 and one study did not report baseline data, although the authors described the study as a
349 pre/post, cluster randomized intervention-control design.[30]

350 Five studies for PA and three studies for SB were included in the meta-analysis as
351 they used a two-group (intervention and control), pre-post design. An interaction effect of the
352 study could be determined and these effects could be pooled together to provide an overall
353 effect of the interventions. The results of the meta-analysis suggest that classroom-based
354 interventions have no significant or meaningful effect on PA or SB in early secondary school
355 adolescents.

356 All interventions were implemented within the classroom and fitted into the school
357 curriculum but none reported incorporating movement into the classroom. Incorporating
358 activity and movement into the learning environment has shown positive effects on PA[44] and
359 can enhance teaching and learning,[45] however most of this evidence derives from primary
360 school settings, indicating the need to expand this research into the secondary school
361 environment.

362 The studies that measured psychological outcomes[29,30,33,35,37] showed overall positive
363 effects on self-efficacy, attitudes, motivation, and knowledge, however in some cases, this
364 did not transfer into changes in PA.[33,35] Although increasing psychological constructs, such as
365 self-efficacy, motivation and attitudes can facilitate behavior change, the small number of
366 studies in the review that measured psychological constructs makes it difficult to determine
367 why changes in behavior did not occur. However, this could be attributed to the intention-
368 behavior gap whereby there is a weak association between intention and behavior.[46] As these
369 constructs were measured by self-report, there could be an element of social desirability
370 bias[47] in that the pupils may have provided answers that they perceived would be desired by
371 the researchers rather than answers true to them.

372 Implementation methods were reported in only five studies.[29-31,36,37] Naylor et al.[19]
373 identified 22 factors that affect implementation, such as time (which included the workload of
374 the teacher, and other requirements), quality of resources, support of the school, teacher and
375 pupil characteristics, pupil behavior, and the schedule of lessons. Authors of two studies
376 stated that to ensure fidelity, materials were supplied and researchers observed a percentage
377 of the lessons, however the level of fidelity was not reported as part of the study.[29,30]
378 Similarly, results of the fidelity element for Tymms et al.[36] were not reported. Whittemore et
379 al.[37] stated that they consulted with teachers prior to the intervention to ensure high
380 implementation and monitored pupil attendance, but again, results of implementation were
381 not reported. Only one study presented results of their monitored implementation.[31] The
382 results of the direct observation indicated high fidelity of the intervention as the classes
383 delivered matched the manual provided. Implementation fidelity is a key component to
384 interventions and the literature suggests that authors who report monitoring implementation
385 of the intervention have greater impacts on the outcome measured.[18] The common outcome
386 variables measured in all four studies were PA and SB, but the results were varied. Therefore
387 it is difficult to determine the impact of implementation factors, such as the ones mentioned
388 by Nayler et al.[19] on the results, especially since the results of the implementation were not
389 reported.

390 One of the quality assessment criteria was study design. Seven studies[29,31,33-37] were
391 rated strong for study design, as they were all randomized cluster control trials. Two studies
392 were rated moderate for design due to their one group pre- and post- cohort design.[29,32] Four
393 of the studies were given a strong rating for confounders.[29,31,36,37] These studies reported
394 controlling for all the primary confounding variables which were applicable to school-based
395 interventions. All studies were rated moderate for the blinding category in the EPHPP. The
396 authors of this review agreed that the pupils may have had knowledge on what the research

was, especially if the intervention classes were in the same school as the control classes, which applied to the studies by Schwarzer et al.[34] and Whittemore et al.[37] The authors added a 'Not Relevant' option for the item related to blinding of assessors. Five studies were rated strong for validity and reliability of the measure used.[29,33,35,36] The strong rating was awarded due to reporting sufficient evidence of measurement validity and reliability. For the participant withdrawal section of the EPHPP, four studies were rated strong for having 90% or more completing the study.[31,35,36,37]

Limitations

Study Limitations

There were a number of limitations at the study and outcome level of this review. Four of the studies reported using convenience sampling to access participants. Although this is rated as weak due to the low likelihood of a true representation of the target population, it should be noted that when researching within the education system, recruitment is often determined by which schools (principals and teachers) support the project proposed[48].

Length of the nine interventions ranged from one 1-hour lesson[34] to 24 lessons over 10 weeks.[30] Not only is this a substantial difference in regards to exposure of the intervention, but some interventions included nutritional elements. Some of these studies reported the number of sessions dedicated to PA/SB[31,32] however some did not.[28,29,35] This makes it difficult when reviewing these studies to determine the true exposure of pupils to the PA/SB elements of the intervention and whether this could have influenced the effectiveness of the interventions.

This review has shown that targeting participants within specific particular Stages of Change could have the greatest positive impact on PA.[34] However, targeting specific Stages of Change strategies in the classroom may be difficult as pupils are already enrolled in the

421 classes and depending on the education system, it may not be feasible to rearrange classes or
422 target particular pupils within an existing class group setting. Very little is reported on
423 variables that could affect implementation.

424 **Review Limitations**

425 This systematic review and meta-analysis has numerous strengths. To our knowledge,
426 this is the first review to summarize and analyze classroom-based PA and SB interventions in
427 secondary/ middle/ high school adolescents aged 11-15 years. However, there are some
428 limitations. This review only included articles that were published in English and did not
429 include other sources (e.g. conference abstracts). Caution should be taken when reviewing
430 the meta-analysis section. Findings from the meta-analysis suggest there was a degree of
431 statistical heterogeneity for both PA and SB. This variance might be attributed to the
432 methodological differences in design and outcome measures of PA and SB for each study
433 included in the meta-analysis. Furthermore, only studies that had a control group were
434 included in the analysis therefore there were no statistical analyses performed on the two pre-
435 and post- cohort studies despite reporting significant results. A number of difficulties arose
436 when performing the meta-analysis. In situations where outcome measures were presented
437 separately (moderate and vigorous PA)[35,37], only moderate PA was included. Schwarzer et
438 al.[34] presented two intervention groups vs. a control. Results of the two intervention groups
439 were combined to form a single intervention group. A meta-analysis was not performed for
440 psychological constructs due to the different constructs being measured i.e. self-efficacy and
441 motivation are different and therefore should not be compared within a meta-analysis.

442 **Conclusion**

443 Overall, there appears to be no clear classroom-based methodology for effectively
444 increasing PA and reducing SB in early secondary school adolescents. This is likely due to

445 the lack of research in this area. The overall findings of this review agree with Russ et al.[12] in
446 that these interventions only produced small effects on PA and SB. The meta-analysis has
447 shown that currently, either classroom-based PA/SB or PA only interventions have no effect
448 on increasing PA or reducing SB, however this evidence is limited due to the lack of studies
449 providing two group, pre- and post- data. There is still little research regarding school-based
450 interventions on reducing SB, and the effectiveness of these interventions is still largely
451 unknown.[49] The results of this review support this statement.

452 The emerging evidence shows there is a positive association between increasing PA,
453 and reducing SB on academic attainment and on-task behavior. Studies suggest that levels of
454 PA decline as children enter secondary school and transition into adolescence.[50]
455 Contradictory to this view, a review has suggested that PA decreases before children enter
456 adolescence.[51] Regardless, physical inactivity is a global issue for adolescents and programs
457 that focus more on all aspects PA and SB, rather than sport and physical education, could
458 help break down barriers and increase motivation and positive attitudes towards PA, and
459 reduce SB, as shown in this review.

460 More research is needed in secondary/middle/ high schools in regards to active
461 classrooms (where movement is incorporated into the learning environment) as little has been
462 done in this age group and setting, and much more rigorous reporting of implementation is
463 vital so that researchers can understand the variables that influence the implementation of
464 such interventions.

465 **Acknowledgements**

466 Thanks are given to Irene Stirling and Dr Stephen Corson (University of Strathclyde),
467 and Dr Anne Martin (University of Edinburgh) for their consultations.

468

469 **References**

[1]Strong WB, Malina RM, Blimkie CJ, Daniels ST, Dishman RK, Gutin B, et al. Evidenced based physical activity for school-age youth. *J Pediatr.* 2005; 146: 732-737.

[2]Biddle SJH, Asare M. Physical activity and mental health in children and adolescents: a review of reviews. *Br J Sports Med.* 2011; 45(11): 886-895. **doi: 10.1136/bjsports-2011-090185**

[3]Currie C, Van der Sluijs W, Whitehead R, Currie D, Rhodes G, Neville F, Inchley J. HBSC 2014 Survey in Scotland National Report. Child and Adolescent Health Research Unit (CAHRU) St. Andrews: 2015. http://www.cahru.org/content/03-publications/04-reports/hbsc_nr14_interactive_final.pdf. Accessed December 13, 2016

[4]Centres for Disease Control and Prevention. Youth physical activity: the role of schools. http://www.cdc.gov/healthyschools/physicalactivity/toolkit/factsheet_pa_guidelines_schools.pdf. Accessed October 13, 2016

[5]World Health Organisation. School policy framework. http://www.who.int/dietphysicalactivity/SPF-en-2008.pdf. Accessed October 13, 2016

[6]Dobbins M, Husson H, DeCorby K, LaRocca RL. School-based physical activity programs for promoting physical activity and fitness in children and adolescents aged 6 to 18. *Cochrane Database Syst Rev.* 2013; (2): **doi: Artn Cd00765110.1002/14651858.Cd007651.Pub2**

[7]Hynynen S, van Stalen MM, Sniehotta FF, Araujo-Soares V, Hardeman W, Chinapaw MJM, et al. (2015). A systematic review of school-based interventions targeting physical activity and sedentary behaviour in older adolescents. *Int Rev Sport Exerc Psychol.* 2015; 9(1): 22-44.

[8]Kriemler S, Meyer U, Martin E, van Sluijs EMF, Andersen LB, Martin BW. Effect of school-based interventions on physical activity and fitness in children and

adolescents: a review of reviews and systematic update. *Br J of Sports Med.* 2011; 45(11): 923-930. **doi: 10.1136/bjsports-2011-090186**

[9]Norris E, Shelton N, Dunsmuir S, Duke-Williams O, Stamatakis E. Physically active lessons as physical activity and educational interventions: A systematic review of methods and results. *Prev Med*. 2015; 72C: 116-125. **doi: 10.1016/j.ypmed.2014.12.027**

[10]Pardo BM, Bengoechea EG, Lanaspa EG, Bush PL, Casterad JZ, Clemente JAJ, Gonzalez LG. (2013). Promising school-based strategies and intervention guidelines to increase physical activity of adolescents. *Health Educ Res.* 2013; 28(3): 523-538. **doi: 10.1093/her/cyt040**

[11]Rafferty R, Breslin G, Brennan D, Hassan D. A systematic review of school-based physical activity interventions on children's wellbeing. *Int Rev Sport Exerc Psychol.* 2016; 9(1): 215-230.

[12]Russ L, Webster CA, Beets MW, Phillips DS. (2015). Systematic Review and Meta-Analysis of School-Based Multicomponent Physical Activity Interventions. *Res Q Exerc Sport.* 2015; 86: A85-A86.

[13]van Sluijs EMF, McMinn AM, Griffin S. Effectiveness of interventions to promote physical activity in children and adolescents: systematic review of controlled trials. *BMJ.* 2007; 335(7622): 703-707. **doi: 10.1136/bmj.39320.843947.BE**

[14]Mahar MT, Murphy SK, Rowe DA, Golden J, Shields AT, Raedeke TD. Effects of a classroom-based program on physical activity and on-task behavior. *Med. Sci. Sports Exerc.* 2006; 38(12): 2086-2094.

[15]Corder K, Atkin AJ, Bamber DJ, et al. Revising on the run or studying on the sofa: prospective associations between physical activity, sedentary behaviour, and exam results in British adolescents. *Int J Behav Nutr Phys Act*. 2015;12: 106-113.

518 [16]Trudeau, F., & Shephard, R.J. (2008) Physical education, school physical activity, school
519 sports and academic performance. *Int J Behav Nutr Phys Act.* 2008; 5: 10-21.
520 **doi:10.1186/1479-5868-5-10**

521 [17]Glasgow RE, Vogt TM, Boles SM. Evaluating the public health impact of health promotion
522 interventions: The RE-AIM framework. *Am. J. Public Health*. 1999; 89(9): 1322-
523 1327

524 [18]Durlak JA, DuPre EP. Implementation matters: A review of research on the influence of
525 implementation on program outcomes and the factors affecting implementation. *Am*
526 *J Community Psychol.* 2008; 41(3-4): 327-350. **doi: 10.1007/s10464-008-9165-0**

527 [19]Naylor PJ, Nettlefold L, Race D, Hoy C, Ashe MC, Higgins JW, McKay HA (2015).
528 Implementation of school based physical activity interventions: A systematic review.
529 *Prev Med.* 2015; 72: 95-115. **doi: 10.1016/j.ypmed.2014.12.034**

530 [20]Helgeson Jr JL. The impact of phyical activity on academics in English classes at the junior
531 high school level. *Diss Abstr Int Sect A Humanit Soc Sci*. 2014; 74.

532 [21]Martins J, Marques A, Peralta M, Palmeira A, da Costa FC. Correlates of physical activity
533 in young people: A narrative review of reviews. Implications for physical education
534 based on a socio-ecological approach. *Retos*. 2017; 31: 292-299

535 [22]Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JPA, et al. The
536 PRISMA Statement for Reporting Systematic Reviews and Meta-Analyses of Studies
537 That Evaluate Health Care Interventions: Explanation and Elaboration. *PLoS Med.*
538 2009; 6(7): e1000100. **doi:10.1371/journal.pmed.1000100**

539 [23]Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group. Preferred Reporting
540 Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *PLoS*
541 *Me.* 2009; 6(7): e1000097. **doi:10.1371/journal.pmed.1000097**

[24]Effective Public Health Practice Project. Quality assessment tool for quantitative studies. http://www.ephpp.ca/tools.html. Accessed December 12, 2016.

[25]Chillon P, Evenson KR, Vaughn A, Ward DS. A systematic review of interventions for promoting active transportation to school. *Int J Behav Nutr Phys Act*. 2011; 14(8): 10.

[26]Higgins, J.P.T., & Green, S. Cochrane Handbook for Systematic Reviews of Interventions (version 5.1.0). The Cochrane Collaboration. 2011. www.cochrane-handbook.org. Accessed October 13, 2016

[27]Cohen, J. *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). Hillsdale, N.J.: Lawrence Erlbaum; 1988.

[28]The Cochrane Collaboration. Review Manager (RevMan) (Version 5.3) [computer software]. Copenhagen: The Nordic Cochrane Centre; 2014.

[29]Contento IR, Koch PA, Lee H, Sauberli W, Calabrese-Barton A. Enhancing personal agency and competence in eating and moving: formative evaluation of a middle school curriculum—Choice, Control, and Change. *J Nutr Educ Behav.* 2007; 39: S179-S186.

[30]Contento IR, Koch PA, Lee H, Calabrese-Barton A. Adolescents demonstrate improvement in obesity risk behaviors following completion of Choice, Control, and Change, a curriculum addressing personal agency and autonomous motivation. *J Am Diet Assoc*. 2010; 110(12): 1830-1839.

[31]Cui, Z., Shah, S., Yan, L., et al. Effect of a school-based peer education intervention on physical activity and sedentary behaviour in Chinese adolescents: A pilot study. *BMJ Open.* 2012; 2(3).

[32]Dunton GF, Lagloire R, Robertson T. Using the RE-AIM Framework to Evaluate the Statewide Dissemination of a School-Based Physical Activity and Nutrition

566 Curriculum: "Exercise Your Options". *Am J Health Promot*. 2009; 23(4):229-232.
567 **doi: 10.4278/ajhp.071211129**
568 [33]Ghaffari M, Sharifirad G, Malekmakan E, Hassanzadeh A. Effect of educational
569 intervention on physical activity-related knowledge, attitude and behavior of among
570 first-grade students of male high schools. *J Educ Health Promot*. 2013; 2: 4-4. **doi:**
571 **10.4103/2277-9531.106642**
572 [34]Schwarzer, R., Cao, D. S., & Lippke, S. (2010). Stage-matched minimal interventions to
573 enhance physical activity in Chinese adolescents. *J Adolesc Health*. 2010; 47(6): 533-
574 539. **doi: 10.1016/j.jadohealth.2010.03.015**
575 [35]Spruijt-Metz D, Nguyen-Michel ST, Goran MI, Chou CP, Huang TTK. (2008). Reducing
576 sedentary behavior in minority girls via a theory-based, tailored classroom media
577 intervention. *Int J Pediatr Obes*. 2008; 3(4): 240-248. **doi:**
578 **10.1080/17477160802113415**
579 [36] Tymms PB, Curtis SE, Routen AC. Clustered randomised controlled trial of two education
580 interventions designed to increase physical activity and well-being of secondary
581 school students: the MOVE Project. *BMJ Open.* 2016; 6.
582 [37]Whittemore R, Jeon S, Grey M. An internet obesity prevention program for adolescents. *J*
583 *Adolesc Health.* 2013; 52(4): 439-447.
584 [38]Deci EL, Ryan RM. *Intrinsic motivation and self-determination in human behavior*. New
585 York, NY: Plenum; 1985.
586 [39]Bandura A. *Social foundations of thought and action: A social cognitive theory*. Englewood
587 Cliffs, NJ: Prentice- Hall, Inc; 1986.
588 [40]Ajzen I. The theory of planned behavior. *Organ Behav Hum Decis Process*. 1991; 50(2):
589 179-211.

590 **1.** [41]Bandura A. *Social Learning Theory*. New York: General Learning Press; 1977

591 [42]Spruijt-Metz, D. *Adolescence, Affect and Health.* London: Psychology Press; 1999

592 [43]Prochaska J, DiClemente C. Stages and processes of self-change in smoking: toward an
593 integrative model of change. *J Consul Clin Psychol*. 1983; 5: 390-395.

594 [44]Martin R, Murtagh EM. (2015). Preliminary findings of Active Classrooms: an intervention
595 to increase physical activity levels in primary school children. *Teach Teach Educ*.
596 2015; 52: 113-127.

597 [45]Norris E, Shelton N, Dunsmuir S, Duke-Williams O, Stamatakis E. Virtual field trips as
598 physically active lessons for children: a pilot study. *BMC Public Health.* 2015; 15:
599 366.

600 [46]Rhodes RE, Dickau L. Moderators of the intention-behaviour relationship in the physical
601 activity domain: a systematic review. *Br J Sports Med.* 2012; 47: 215-225.
602 **doi:10.1136/bjsports-2011-090411**

603 [47]Grimm P. Social desirability bias. *Wiley International Encyclopedia of Marketing*. John
604 Wiley & Sons, Ltd; 2010.

605 [48]Rice M, Bunker KD, Kang DH, Howell CC, Weaver M. Assessing and recruiting children
606 for research in schools. *Western J Nurs Res*. 2007; 29(4): 501-514.

607 [49]Hegarty LM, Mair JL, Kirby K, Murtagh E, Murphy MH. School-based interventions to
608 reduce sedentary behaviour in children: a systematic review. *AIMS Public Health.*
609 2016; 3(3): 520-541.

610 [50]Marks J, Barnett LM, Strugnell C, Allender S. Changing from primary to secondary school
611 highlights opportunities for school environment interventions aiming to increase
612 physical activity and reduce sedentary behaviour: a longitudinal cohort study. *Int J*
613 *Behav Nutr Phys Act*. 2015; 12: 59. **doi: 10.1186/s12966-015-0218-0**

[51]Reilly JJ. When does it all go wrong? Longitudinal studies of changes in moderat-to-vigorous-intensity physical activity across childhood and adolescence. *J Exerc Sci Fit. 2016; 14(1): 1-6.*

Tables

Table 1. Search strategy used to retrieve potential articles

Population	(Adolescen* **OR** teenage* **OR** youth **OR** pupils **OR** child*)
Search operator	**AND**
Setting	(school* **OR** class* **OR** lesson*)
Search operator	**AND**
Outcome	(Physical activity **OR** walk* **OR** mov* **OR** activity breaks **OR** exercise* **OR** stand*)
Search operator	**AND**
Outcome	(Sedentary behaviour **OR** sedentary behavior **OR** sitting time **OR** sit*)
Search operator	**AND**
Study design	(Interventions **OR** randomised controlled trial **OR** randomized controlled trial **OR** (pre and post) **OR** quasi experimental)
Search operator	**NOT**

Exclusion	(Physical education)

619

Table 2. Summary of studies included in the review.

Author	Country	Design	Population	Intervention	Results
Cui et al. (2012)	China	CRCT	12.7 years N = 738 Mixed gender	Peer-educational intervention (peers teaching educational content to those in their year on PA and SB)	No significant increases compared to control. MVPA (min/day) ($p = 0.83$, $d = 0.02$). MVPA in school (min/day) ($p = 0.52$, $d = -0.026$) post intervention. No significant difference in sedentary behaviours ($p = 0.21$, $d = -0.025$) post intervention. Only significant reduction in sedentary behaviour was on weekdays & computer usage ($p<0.05$) at 7 month follow up.
Ghaffari et al. (2013)	Iran	CRCT	14.0 years N = 85 Boys only 1st grade of High School	Educational intervention	Significant increase and large effect on knowledge & attitude scores for intervention group ($p < 0.001$, $d = 1.94$ and 1.71, respectively) at time point 2. No significant difference ($p = 0.390$) ($d = -0.38$) in PA.
Spruijt-Metz et al. (2008)	USA	CRCT	12.5 years N = 459 Girls only	Classroom media intervention	No significant differences ($p > 0.05$) – Light activity ($d = 0.043$), Moderate activity ($d = -0.07$), high activity ($d = 0.04$). TV/ video game/ internet significantly decreased ($p < 0.05$, $d = -0.28$). Significant increase in intrinsic motivation ($p < 0.05$, $d = 0.11$)
[a]Dunton et al. (2009)	USA	Cohort (pre and post design)	12.47 years N = 695 Mixed gender	"Exercise Your Options"	Significant increase in PA ($p < 0.001$, $d = 0.2$) & significant decrease in video games (hours/day) ($p = 0.002$, $d = -0.21$) and TV viewing (hours/day) ($p =0.024$, $d = -0.15$).
[b]Tymms et al. (2016)	UK	CRCT	11-14 years N = 1391 Mixed Gender	Peer mentoring – Year 9 pupil mentors a Year 7 pupil, once a week for six weeks, to work through a booklet to help promote and increase PA. Participative Learning – Six lessons in Geography which uses GPS to allow Year 7 pupils to collect data on their own PA.	No significant differences on daily MVPA between the Peer-Mentoring ($p > 0.05$, $d = -0.01$), Participative Learning ($d = 0.36$), or a combination of both ($d = -0.02$) compared to the control.

Author	Country	Design	Population	Intervention	Results
Schwarzer et al. (2010)	China	CRCT	13.8 years N = 534 Mixed gender	Resource communication (emphasizing the importance of PA and discussing the pros and cons) Planning intervention (ways to overcome barriers to PA).	Significant increases in PA were reported between pre- and post- resource communication ($p < 0.01$, $d = 0.96$) for preintenders. This increase was non-significant in intenders and actors ($p > 0.05$, $d = 0.08$, $d = 0.01$). There were no significant increases in PA in the planning intervention ($p > 0.05$, $d = 0.22$).
[b]Contento et al. (2010)	USA	CRCT	12.0 years N = 1136 Mixed gender	"Choice, Control and Change" an educational intervention delivered in science/ physical education class (classroom-based)	Significant increase in walking for transportation ($p < 0.001$, $d = 0.26$), walking for exercise (p 0.044, $d = 0.14$), stairs for exercise ($p < 0.001$, $d = 0.26$). Leisure screen time significantly decreased ($p < 0.001$, $d = -0.38$). Significant increase in competence and autonomy ($p = 0.005$, $d = 0.22$, $d = 0.24$, respectively). Significant increase in self-efficacy ($p < 0.001$, $d = 0.37$) and intentions to do more PA ($p = 0.012$, $d = 0.18$).
[a]Contento et al. (2007)	USA	Cohort (pre and post design)	12.0 years N = 278 Mixed gender	"Choice, Control and Change" an educational intervention delivered in science/ physical education class (classroom-based)	No significant difference for walking ($p = 0.830$, $d = 0.02$) or stair use ($p = 0.867$, $d = 0.01$). Significantly decreased days/week playing video games ($p < 0.001$, $d = -0.25$), scores for minutes per day ($p < 0.001$, $d = -0.27$). TV viewing days/week ($p = 0.003$, $d = -0.18$), scores for minutes per day TV viewing ($p <0.001$, $d = -0.3$).
[c]Whittemore et al. (2013)	USA	CRCT	15.3 years N = 384 Mixed gender	HEALTH[e]TEEN HEALTH[e]TEEN + Coping Skills Training (CST)	No significant differences between groups for moderate or vigorous PA ($p > 0.05$, $d = 0.18$), SB (weekdays or weekends) ($p > 0.05$, $d = -0.09$, $d = 0.04$) or self-efficacy ($p > 0.05$, $d = 0.08$).

***** Effect sizes presented are interaction effect sizes.
[a] = cohort pre- and post- design study. Effect sizes presented are for pre- and post- intervention. Not an interaction effect.
[b] = only data for one time point was presented. Effect sizes presented are for the one time point. Not an interaction effect.
[c] = both intervention and control were physical activity promotion programs yet one had additional coping skills training.

Table 3. Summary of ratings for each study under the different elements of the EPHPP tool.[1]

	Selection Bias	Study Design	Confounders	Blinding	Data Collection	Withdrawals and Dropout	Global Rating
Cui et al. (2012)	Weak	Strong	Strong	Moderate	Strong	Strong	Moderate
Spruijt-Metz et al. (2008)	Strong	Strong	Moderate	Moderate	Strong	Strong	Strong
Ghaffari et al. (2013)	Moderate	Strong	Weak	Moderate	Strong	Weak	Weak
Dunton et al. (2009)	Weak	Moderate	Moderate	Moderate	Weak	Moderate	Weak
Whittemore et al. (2013)	Weak	Strong	Strong	Moderate	Moderate	Strong	Moderate
Schwarzer et al. (2010)	Moderate	Strong	Weak	Moderate	Weak	Moderate	Weak
Contento et al. (2010)	Weak	Strong	Strong	Moderate	Strong	Moderate	Moderate
Contento et al. (2007)	Moderate	Moderate	Moderate	Moderate	Moderate	Weak	Moderate
Tymms et al. (2016)	Weak	Strong	Strong	Moderate	Strong	Strong	Moderate

[1] Adapted version of the EPHPP tool was used. For study design, cluster randomised was added and given a strong rating. For blinding, "not relevant" was added as an option.