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\(_{2\text{max}}\) (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. According to Durlak and Dupre, implementation includes fidelity, dosage, quality, participant responsiveness, program differentiation, monitoring of controls, program reach and adaptation. 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. Naylor et al. 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.

Several systematic reviews have examined school-based interventions, of which only one has focused on classroom-based PA interventions. Only one of the original studies included in that review was based in a secondary school, 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), 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.\textsuperscript{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 (Microsofost Corp, Redmond, WA) spreadsheet. Manual inspection of duplicates was performed again to ensure that there were no duplicates.
The inclusion criteria for screening articles were: a) randomized controlled trials (RCTs), controlled trials (CTs), quasi-experimental, or pre- and post- study designs; b) studies based in a classroom setting only, which targeted PA or SB, or both PA and SB; c) non-clinical secondary/ middle/ high school adolescents between the ages of 11-15 years old.

We excluded any study based in primary/elementary schools or in high/secondary schools where interventions targeted adolescents over the age of 15 years. This is because primary and secondary school environments are different in terms of education and the structure of the school day.

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

**Data Collection Process and Data Items**

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

**Risk of Bias in Individual Studies**

An adapted version of the Effective Public Health Practice Project (EPHPP) tool was used for quality assessment. The EPHPP has a rating scale of 1 to 3 (1 = strong, 2 =
classroom activity in adolescents

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

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

### Meta-Analysis

A meta-analysis was performed to determine the overall effect of classroom-based interventions on PA and SB. Review Manager, version 5.3 (RevMan 5.3) (The Nordic Cochrane Centre, Copenhagen) computer software was used to conduct the meta-analysis. Only studies that used a two group (intervention/control), pre-post design were included in the meta-analysis (n = 5). Data from baseline measures and the first measurement post-intervention were analyzed. The data inputted into RevMan 5.3 were: the standardized mean differences between pre- and post-intervention for the intervention group, and the control group; the pooled SD of the four cells of data (pre- and post-intervention data, pre- and post-control data); and the sample size of each group (n). This produced a Cohen’s d for the interaction effect and 95% confidence interval (CI). The RevMan 5.3 software then pooled the effects for all studies to produce an overall effect, weighted by individual study sample size. Standardized means were calculated to take into account that each study used different 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, and five included for the meta-analysis. 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, two were based in China, one was
based in the UK and one was based in Iran. Seven studies were cluster randomized control
trials and two were pre- and post- cohort design with no control group. Sample
size ranged from N = 85 to N = 1391. The reported mean age of participants ranged from
12.0 years to 15.3 years. Whittermore et al. 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.

Interventions

The interventions included in the study were all classroom-based and were
educational. Three of the studies investigated PA as an outcome and one study
investigated both PA and SB as outcomes. Five studies had a nutritional element to the
program alongside PA and SB. No studies investigated reducing SB only. The nutritional elements included, for example, education on fruit and vegetable consumption and measuring these outcome variables. Five studies measured psychological outcomes including self-efficacy, motivation, and attitudes. Details on the interventions are presented in Table 2.

Theoretical Underpinnings

Six of the eight studies reported using one or more theoretical frameworks to inform their interventions. Two used Self-Determination Theory; two used Social Cognitive Theory; one used the Theory of Planned Behavior; one used Social Learning Theory; one used Theory of Meanings of Behaviour, and one study used a version of Stages of Change model.

Physical Activity/ Sedentary Behavior

All PA data were collected through self-reported measures, except for one study in which PA was measured objectively. Varieties of PA outcome measures were reported. These included: MVPA (mins/day); PA performance; PA expressed as the number of 30-minute blocks spent in each of three intensities (high, medium, light); PA (days/week); moderate exercise (days/week for at least 30 minutes); vigorous exercise (day/week for 20 minutes); PA frequency (presented as a score of 1-4; 1 = never, 2 = 2 times per week, 3 = 3-4 times per week, 4 = almost every day); walking; and stair climbing. There were varied results regarding the effects of the interventions on PA behavior, with only three studies reporting significant results. Contento et al. found a significant increase in walking for transport and walking for exercise (0.55 days/week, p <0.001, d = 0.26; 0.36 days/week, p = 0.044, d = 0.14, respectively) compared to control post intervention. Dunton et al. reported 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. 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 measured outcomes of SB, using self-report. One study measured SB (mins/day); one measured SB (hours/day); one measured screen time (television viewing/game play/internet usage) in hours/day; one measured screen time in half hour blocks; two studies measured screen time in days/week. Four studies reported significant decreases in SB. Dunton et al. 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. 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. 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. reported a significant decrease in screen time in the intervention group compared to the control (p < 0.05, d = -0.28).

Whittemore et al. 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 =
Classroom Activity in Adolescents

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

**Psychological Outcomes**

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

Motivation was measured in two studies. Contento et al. measured pupils’ autonomous motivation and reported significantly (p = 0.005) higher autonomy and competence towards PA in the intervention groups compared to the control group (autonomy = 3.13 ± 0.74 vs 2.94 ± 0.82, d = 0.24; competence = 3.13 ± 0.77 vs 2.95 ± 0.88, respectively, d = 0.22). Spruijt-Metz et al. measured the different constructs of motivation (external regulation, introjected regulation, identified regulation, and intrinsic motivation), with the 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 and walking. Ghaffari 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. 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, five were rated as moderate and three were rated as weak. A summary of the ratings for each category is presented in Table 3.

Implementation

Five studies reported monitoring of implementation. To ensure fidelity, Contento et al. 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. had two members of the research team attend one third of lessons taught by each teacher, provided guidance on how the lessons
should be run and provided all materials. Whittemore et al.\textsuperscript{37} consulted with teachers prior to the intervention to try to optimize implementation. Pupil participation was also monitored by the research team bi-monthly.

Cui et al.\textsuperscript{31} monitored implementation through direct observation. A research member and an external figure observed the peer education classes. Immediately post intervention, a focus group was conducted with pupils and interviews were conducted with staff members of the schools. Results of the observation suggested that the material and classes delivered by the peer leaders met the content and objectives that were presented in the peer leaders’ manual. The data collected through interviews with staff members indicated that the intervention was feasible and acceptable.

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

\textbf{Meta-Analysis}

The results of the random-effects meta-analysis showed there were no significant individual study effects on PA or SB for the interventions included in the analysis. For PA, the overall effect of the interventions across the five included studies was non-significant (p = 0.55, d = 0.05, 95\% CI [-0.11, 0.21]). For SB, the overall effect of the interventions across the three included studies was non-significant (p = 0.16, d = -0.11, 95\% CI [-0.25, 0.04]). The I\textsuperscript{2} for both meta-analyses indicated that there was substantial heterogeneity of the studies (67\% and 52\% for PA and SB respectively). The I\textsuperscript{2} percentage determines the variance that could be attributed to the heterogeneity of the studies included for analysis. Forest plots of 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, these effects were only small, and were based on post-intervention data only and single group, pre- and post- intervention data. The study by Schwarzer et al. 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). 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 but these were only small effects. Importantly, two of these studies lacked a control group.
and one study did not report baseline data, although the authors described the study as a pre/post, cluster randomized intervention-control design.\textsuperscript{30}

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

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

The studies that measured psychological outcomes\textsuperscript{29,30,33,35,37} showed overall positive effects on self-efficacy, attitudes, motivation, and knowledge, however in some cases, this did not transfer into changes in PA.\textsuperscript{33,35} Although increasing psychological constructs, such as self-efficacy, motivation and attitudes can facilitate behavior change, the small number of studies in the review that measured psychological constructs makes it difficult to determine why changes in behavior did not occur. However, this could be attributed to the intention-behavior gap whereby there is a weak association between intention and behavior.\textsuperscript{46} As these constructs were measured by self-report, there could be an element of social desirability bias\textsuperscript{47} in that the pupils may have provided answers that they perceived would be desired by the researchers rather than answers true to them.
Implementation methods were reported in only five studies. Naylor et al. identified 22 factors that affect implementation, such as time (which included the workload of the teacher, and other requirements), quality of resources, support of the school, teacher and pupil characteristics, pupil behavior, and the schedule of lessons. Authors of two studies stated that to ensure fidelity, materials were supplied and researchers observed a percentage of the lessons, however the level of fidelity was not reported as part of the study. Similarly, results of the fidelity element for Tymms et al. were not reported. Whittemore et al. stated that they consulted with teachers prior to the intervention to ensure high implementation and monitored pupil attendance, but again, results of implementation were not reported. Only one study presented results of their monitored implementation. The results of the direct observation indicated high fidelity of the intervention as the classes delivered matched the manual provided. Implementation fidelity is a key component to interventions and the literature suggests that authors who report monitoring implementation of the intervention have greater impacts on the outcome measured. The common outcome variables measured in all four studies were PA and SB, but the results were varied. Therefore it is difficult to determine the impact of implementation factors, such as the ones mentioned by Naylor et al. on the results, especially since the results of the implementation were not reported.

One of the quality assessment criteria was study design. Seven studies were rated strong for study design, as they were all randomized cluster control trials. Two studies were rated moderate for design due to their one group pre- and post- cohort design. Four of the studies were given a strong rating for confounders. These studies reported controlling for all the primary confounding variables which were applicable to school-based interventions. All studies were rated moderate for the blinding category in the EPHPP. The 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.\textsuperscript{34} and Whittemore et al.\textsuperscript{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.\textsuperscript{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.\textsuperscript{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\textsuperscript{48}.

Length of the nine interventions ranged from one 1-hour lesson\textsuperscript{34} to 24 lessons over 10 weeks.\textsuperscript{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\textsuperscript{31,32} however some did not.\textsuperscript{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.\textsuperscript{34} However, targeting specific Stages of Change strategies in the classroom may be difficult as pupils are already enrolled in the
classes and depending on the education system, it may not be feasible to rearrange classes or
target particular pupils within an existing class group setting. Very little is reported on
variables that could affect implementation.

Review Limitations

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

Conclusion

Overall, there appears to be no clear classroom-based methodology for effectively
increasing PA and reducing SB in early secondary school adolescents. This is likely due to
the lack of research in this area. The overall findings of this review agree with Russ et al.\textsuperscript{12} in that these interventions only produced small effects on PA and SB. The meta-analysis has shown that currently, either classroom-based PA/SB or PA only interventions have no effect on increasing PA or reducing SB, however this evidence is limited due to the lack of studies providing two group, pre- and post- data. There is still little research regarding school-based interventions on reducing SB, and the effectiveness of these interventions is still largely unknown.\textsuperscript{49} The results of this review support this statement.

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

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

\textbf{Acknowledgements}

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

\textbf{References}


Classroom Activity in Adolescents


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

doi: 10.4278/ajhp.0712111129


10.4103/2277-9531.106642


10.1080/17477160802113415


Classroom Activity in Adolescents


### Tables

**Table 1. Search strategy used to retrieve potential articles**

| Population | (Adolescent* 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 move* 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)
Table 2. Summary of studies included in the review.

<table>
<thead>
<tr>
<th>Author</th>
<th>Country</th>
<th>Design</th>
<th>Population</th>
<th>Intervention</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cui et al. (2012)</td>
<td>China</td>
<td>CRCT</td>
<td>12.7 years N = 738 Mixed gender</td>
<td>Peer-educational intervention (peers teaching educational content to those in their year on PA and SB)</td>
<td>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 &amp; computer usage (p&lt;0.05) at 7 month follow up.</td>
</tr>
<tr>
<td>Ghaffari et al.</td>
<td>Iran</td>
<td>CRCT</td>
<td>14.0 years N = 85 Boys only 1st grade of High School</td>
<td>Educational intervention</td>
<td>Significant increase and large effect on knowledge &amp; attitude scores for intervention group (p &lt; 0.001, d = 1.94 and 1.71, respectively) at time point 2. No significant difference (p = 0.390) (d = -0.38) in PA.</td>
</tr>
<tr>
<td>Spruijt-Metz et al. (2008)</td>
<td>USA</td>
<td>CRCT</td>
<td>12.5 years N = 459 Girls only</td>
<td>Classroom media intervention</td>
<td>No significant differences (p &gt; 0.05) – Light activity (d = 0.043), Moderate activity (d = -0.07), high activity (d = 0.04). TV/ video game/ internet significantly decreased (p &lt; 0.05, d = -0.28). Significant increase in intrinsic motivation (p &lt; 0.05, d = 0.11)</td>
</tr>
<tr>
<td>Dunton et al.</td>
<td>USA</td>
<td>Cohort (pre and post design)</td>
<td>12.47 years N = 695 Mixed gender</td>
<td>“Exercise Your Options”</td>
<td>Significant increase in PA (p &lt; 0.001, d = 0.2) &amp; significant decrease in video games (hours/day) (p = 0.002, d = -0.21) and TV viewing (hours/day) (p =0.024, d = -0.15).</td>
</tr>
<tr>
<td>Tymms et al.</td>
<td>UK</td>
<td>CRCT</td>
<td>11-14 years N = 1391 Mixed Gender</td>
<td>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.</td>
<td>No significant differences on daily MVPA between the Peer-Mentoring (p &gt; 0.05, d = -0.01), Participative Learning (d = 0.36), or a combination of both (d = -0.02) compared to the control.</td>
</tr>
<tr>
<td>Author</td>
<td>Country</td>
<td>Design</td>
<td>Population</td>
<td>Intervention</td>
<td>Results</td>
</tr>
<tr>
<td>-------------------</td>
<td>---------</td>
<td>------------------</td>
<td>------------------</td>
<td>-----------------------------------------------------------------------------</td>
<td>---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Schwarzer et al.</td>
<td>China</td>
<td>CRCT</td>
<td>13.8 years</td>
<td>Resource communication (emphasizing the importance of PA and discussing the pros and cons) Planning intervention (ways to overcome barriers to PA).</td>
<td>Significant increases in PA were reported between pre- and post- resource communication (p &lt; 0.01, d = 0.96) for pretenders. This increase was non-significant in intenders and actors (p &gt; 0.05, d = 0.08, d = 0.01). There were no significant increases in PA in the planning intervention (p &gt; 0.05, d = 0.22).</td>
</tr>
<tr>
<td>Contento et al.</td>
<td>USA</td>
<td>CRCT</td>
<td>12.0 years</td>
<td>“Choice, Control and Change” an educational intervention delivered in science/ physical education class (classroom-based)</td>
<td>Significant increase in walking for transportation (p &lt; 0.001, d = 0.26), walking for exercise (p 0.044, d = 0.14), stairs for exercise (p &lt; 0.001, d = 0.26). Leisure screen time significantly decreased (p &lt; 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 &lt; 0.001, d = 0.37) and intentions to do more PA (p = 0.012, d = 0.18).</td>
</tr>
<tr>
<td>Contento et al.</td>
<td>USA</td>
<td>Cohort (pre and post design)</td>
<td>12.0 years</td>
<td>“Choice, Control and Change” an educational intervention delivered in science/ physical education class (classroom-based)</td>
<td>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 &lt; 0.001, d = -0.25), scores for minutes per day (p &lt; 0.001, d = -0.27). TV viewing days/week (p = 0.003, d = -0.18), scores for minutes per day TV viewing (p &lt;0.001, d = -0.3).</td>
</tr>
<tr>
<td>Whittemore et al.</td>
<td>USA</td>
<td>CRCT</td>
<td>15.3 years</td>
<td>HEALTH[e]TEEN HEALTH[e]TEEN + Coping Skills Training (CST)</td>
<td>No significant differences between groups for moderate or vigorous PA (p &gt; 0.05, d = 0.18), SB (weekdays or weekends) (p &gt; 0.05, d = -0.09, d = 0.04) or self-efficacy (p &gt; 0.05, d = 0.08).</td>
</tr>
</tbody>
</table>

* 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.\(^1\)

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

\(^1\) 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.