A systematic review of school-based interventions targeting physical activity and sedentary behaviour among older adolescents


ABSTRACT
Lack of physical activity (PA) and high levels of sedentary behaviour (SB) have been associated with health problems. This systematic review evaluates the effectiveness of school-based interventions to increase PA and decrease SB among 15–19-year-old adolescents, and examines whether intervention characteristics (intervention length, delivery mode and intervention provider) and intervention content (i.e. behaviour change techniques, BCTs) are related to intervention effectiveness. A systematic search of randomised or cluster randomised controlled trials with outcome measures of PA and/or SB rendered 10 results. Risk of bias was assessed using the Cochrane risk of bias tool. Intervention content was coded using Behaviour Change Technique Taxonomy v1. Seven out of 10 studies reported significant increases in PA. Effects were generally small and short-term (Cohen’s $d$ ranged from 0.132 to 0.659). Two out of four studies that measured SB reported significant reductions in SB. Interventions that increased PA included a higher number of BCTs, specific BCTs (e.g., goal setting, action planning and self-monitoring), and were delivered by research staff. Intervention length and mode of delivery were unrelated to effectiveness. More studies are needed that evaluate long-term intervention effectiveness and target SBs among older adolescents.

ARTICLE HISTORY
Received 22 November 2014
Accepted 10 July 2015

KEYWORDS
physical activity; sedentary behaviour; adolescents; school-based intervention; behaviour change techniques

Introduction
Ample evidence exists of the profound benefits of leading a physically active life. Regular physical activity (PA) has been associated with a decreased risk of physical and mental health problems, such as obesity, cardiovascular disease, type II diabetes and depression.
among children and adolescents (Biddle, Gorely, & Stensel, 2004; Janssen & LeBlanc, 2010; Penedo & Dahn, 2005). The evidence on the adverse health effects of excessive sedentary behaviour (SB) in youth is inconsistent to date (Chinapaw, Altenburg, & Brug, 2015; Chinapaw, Proper, Brug, van Mechelen, & Singh, 2011). Yet there is an indication that excessive SB is associated with both physical and psychological health problems such as unfavourable body composition, decreased fitness, lowered self-esteem and pro-social behaviour and decreased academic achievement in school-aged children and youths (Tremblay et al., 2011a). If non-sedentary lifestyle habits can be initiated in adolescence, this may have beneficial preventive value, as the effects on mortality and morbidity among adults have already been established (Thorp, Owen, Neuhaus, & Dunstan, 2011; Wilmot et al., 2012). To achieve and maintain good health, children and adolescents are recommended to engage in at least 60 minutes of moderate- to vigorous-intensity physical activity (MVPA) each day (Janssen & LeBlanc, 2010; Strong et al., 2005). However, according to the Health Behaviour in School-aged Children study 2009–2010, only 15% of 15-year-old adolescents met this recommendation globally (Currie et al., 2012). There are also recommendations for sedentary time – although not as widely agreed upon as for PA. For example, the guidelines of the Canadian Society for Exercise Physiology (CSEP) for decreasing SB in youth state that the goal of minimising the time spent sedentary can be achieved by several means: in addition to limiting recreational screen time, also limiting sedentary transport, extended sitting time and time spent indoors throughout the day which also may involve time spent in classrooms (Tremblay et al., 2011b). In Finland recent national recommendations on the reduction of sedentary time explicitly identify schools as one of the key settings (Ministry of Social and Health Affairs, 2015). As PA even decreases during adolescence (Dumith, Gigante, Domingues, & Kohl, 2011) and youths spend a lot of their time both at home and in school being sedentary (van Stralen et al., 2014; Verloigne et al., 2012), interventions that aim to promote PA and to reduce SB among adolescents are urgently needed.

Schools are a promising setting for health promotion interventions aimed at adolescents, since they reach a majority of the target population. In previous studies, school-based interventions have been found to have significant effects on adolescents’ PA and SB. However, the effects have been small, short-term and have largely differed between interventions (Biddle, O’Connell, & Braithwaite, 2011; De Meester, van Lenthe, Spittaels, Lien, & De Bourdaedhuij, 2009; Demetriou & Höner, 2012; Dobbins, De Corby, Robeson, Husson, & Tirilis, 2009; Lonsdale et al., 2013; Metcalf, Henley, & Wilkin, 2012; van Sluijs, McMinn, & Griffin, 2007). Also, effects are mostly seen in school-related PA while effects outside of this context in leisure time and commuting PA are often not observed or assessed (Kriemler et al., 2011). Furthermore, age might moderate the effectiveness of school-based interventions. Earlier reviews have tended to either include a broad age range (e.g., 6–18-year-olds) or focus on younger age categories (Safron, Cislak, Gaspar, & Luszczynska, 2011; van Sluijs et al., 2007). To our knowledge, no recent review has focused solely on interventions targeting older adolescents, and evidence in this age group is lacking. These previous studies have disregarded the possibility that age groups differ in developmental stages and thus require different intervention strategies. Since older adolescents have been a less studied target group, this review focuses on them exclusively.
The heterogeneity in effectiveness of school-based PA interventions (see, e.g., Kriemler et al., 2011; van Sluijs et al., 2007) may be explained by differences in factors such as length of the intervention, mode of delivery, provider or content of the intervention. Yet no studies so far have systematically sought to explain this variability by analysing intervention content (see Abraham & Michie, 2008; Michie et al., 2013). Taxonomies of behaviour change techniques (BCTs) allow more in-depth analyses of interventions and intervention strategies (Abraham & Michie, 2008; Michie et al., 2013).

Three previous reviews have identified the BCTs used in interventions targeting obesity- and weight-related behaviours among children and adolescents (Golley, Hendrie, Slater, & Corsini, 2011; Hendrie et al., 2012; Martin, Chater, & Lorencatto, 2013). In community- and school-based interventions to target obesity-related behaviours in 1–18-year-old children, three BCTs were associated with effectiveness: Providing general information on behaviour – health links, Prompting practice, and Planning for social support/social changes (Hendrie et al., 2012). In their review on interventions that target parents to improve children’s weight-related nutrition intake and activity patterns, Golley et al. (2011) found that effective interventions included: Prompting barrier identification, Restructuring the home environment, Prompting self-monitoring and Prompting specific goal setting. In childhood obesity prevention interventions, only Prompting generalisation of target behaviour was shown to be related to intervention effectiveness (Martin et al., 2013). None of the existing reviews has conducted a BCT analysis of school-based PA or SB interventions among adolescents aged over 15 years old (e.g., Biddle et al., 2011; Dobbins et al., 2009; van Sluijs et al., 2007).

More studies have investigated effective BCTs to increase PA among adult populations (see, e.g., Abraham & Graham-Rowe, 2009; Bird et al., 2013; Michie, Abraham, Whittington, McAteer, & Gupta, 2009; Williams & French, 2011). These systematic reviews show that effective PA interventions have usually included BCTs linked with self-regulation, such as self-monitoring, goal setting and action planning.

There is a lack of evidence about whether the number of BCTs reported as being present in interventions is related to intervention effectiveness. However, a review of interventions targeting obesity and weight-related nutrition and PA behaviours in children both in the home and school/community setting (Hendrie et al., 2012) suggested that interventions including a higher number of BCTs were associated with better outcomes.

In addition to characterising interventions in terms of component BCTs, existing reviews of school-based PA interventions fall short of analysing whether intervention outcomes are influenced by intervention length, mode of delivery (e.g., oral communication, written material, video, interactive computer program, self-help, individual face-to-face, group face-to-face, telephone) or provider of delivery. A systematic review of systematic reviews and meta-analyses on effects of school-based interventions targeting obesity-related behaviours showed that more effective interventions generally lasted at least three months, i.e. a longer duration was reported to be more effective compared to shorter interventions (Safron et al., 2011). A better understanding of whether the above-mentioned characteristics are related to intervention effectiveness would enable evidence-based decisions during intervention development.

Our review provides an up-to-date systematic analysis of the effectiveness of school-based interventions targeting PA and/or SB among older adolescents (aged 15–19 years), using an exploratory narrative review with an analysis of intervention components including BCTs.
**Review questions**

(1) Are school-based interventions to increase PA and/or reduce SB among 15- to 19-year-olds effective?

(2) Does effectiveness of the interventions vary according to a) the length of the intervention, b) the mode of delivery (e.g., oral communication, written material, video, interactive computer program, self-help, individual face-to-face, group face-to-face, telephone), and c) the provider of the intervention (e.g., teacher, peer, researcher)?

(3) Which BCTs have been used in the interventions and how is intervention content related to effectiveness?

**Methods**

The protocol for this systematic review was published in PROSPERO – International prospective register of systematic reviews (Hynynen et al., 2013).

**Study selection and search strategy**

A systematic search was conducted using the following electronic databases of peer-reviewed journal articles and online research registers: Medline, Cinahl, Embase, PsycINFO, Cochrane Central Register of Controlled Trials, Cochrane Database of Systematic Reviews, and Cochrane Methodology Register. The search spanned from the year of the database’s inception up to and including February 2013. During the search we imposed no language restrictions but only articles written in English were included in the review. Individualised search strategies for the different databases included combinations of key words related to PA and SB, school, intervention and adolescence (see supplemental data). We used a modified version of the Cochrane highly sensitive search strategy for identifying randomised trials, where the sections on complementary therapies were removed. The reference lists of studies that met the inclusion criteria as well as other relevant reviews were scanned. The researchers also sent inquiries about ongoing intervention studies via relevant networks, such as scientific societies. However, no additional studies were identified through this approach.

Inclusion criteria specified randomised or cluster randomised studies on school-based interventions that targeted PA and/or SB among 15- to 19-year-old adolescents who are apparently healthy (for a more detailed description, see Hynynen et al., 2013). Included studies could also address other health behaviours, but data on these behaviours were not extracted. Inclusion criteria specified that interventions had to be primarily based in schools but they might also include home- or community-based components. The included studies reported either self-reported or objectively measured (e.g., accelerometer) PA or SB or both, presenting a baseline and a post-intervention measurement. Studies that only targeted individuals at increased health risk (e.g., obese youths) were excluded. We also excluded studies that measured students’ PA during physical education (PE) classes only.

This strategy identified 8470 references that were imported into the RefWorks database (see Figure 1). After excluding duplicates, 5482 references remained. The screening of the remaining references was conducted in three phases by two researchers (STH and MVS).
working independently. Publications were included if eligibility was unclear. In the first phase of screening, titles of all references were reviewed. During this phase 4456 references were excluded. In the second phase, abstracts of the remaining 1026 articles were screened and 667 articles excluded. In the third phase, full text papers were obtained for 359 articles and eligibility reviewed in detail. Any disagreements were resolved by discussion between the two researchers.

We created a standardised form to extract all relevant details of the trial characteristics (study design, number of participants, method of randomisation, study setting and

Figure 1. Identification of the included studies.
country), sample characteristics (inclusion and exclusion criteria for participants, co-morbidities), details of interventions (mode of delivery, provider, setting, recipient, intensity, length), participant characteristics, participant attrition, physical activity and sedentary behaviour outcome measures and possible secondary outcomes. The data extraction was carried out by one researcher (STH) and verified by a second (MVS or another researcher).

**Assessing risk of bias**

The Cochrane risk of bias tool (Higgins et al., 2011) was used for assessing risk of bias in the included articles, covering the allocation procedures, outcome analyses, reporting and other possible sources of bias. Two researchers (STH and MC) independently assessed the quality of all studies that met the inclusion criteria. The following domains were considered: sequence generation, allocation concealment, blinding of study personnel and participants, incomplete outcome data, selective outcome reporting, other possible sources of bias, and when relevant cluster recruitment bias and baseline imbalance bias. Each paper was carefully assessed for each domain and judgements were made regarding potential bias, according to three categories: low risk, high risk, or risk unclear. Inter-rater reliability was calculated (average percentage agreement 74%).

**Coding the behaviour change techniques**

The BCT coding was conducted by two researchers (STH and MVS) using Michie et al.’s (2013) behaviour change technique taxonomy (BCTT v1). Prior to coding, STH had attended a BCT training workshop and MVS completed a group tutorial training programme guided by phone. BCTs were coded separately for increasing PA and reducing SB. Both intervention arms and control arms were coded for BCTs. All published materials (e.g., intervention protocols) of the included trials were used to characterise intervention content. Prior to extracting BCTs from included intervention studies, the two coders piloted consistency in their BCT coding by coding three intervention reports that were not included in the review as they targeted a different age group. All included interventions were then coded independently. Inter-rater reliability was assessed and percent agreement on BCTs present in the descriptions was high (83.3%). Discrepancies between both primary coders were discussed and resolved with a third researcher (WH) who is a member of the BCTT v1 team.

**Assessing intervention and BCT effectiveness**

For each study, intervention effects (Cohen’s $d$) for PA and SB were calculated at the first follow-up post-intervention. This was the only outcome measurement point after baseline that was reported in all included studies. Therefore, we selected this point in time to compare the effects across studies. For the purpose of calculating effect sizes, we contacted four authors for further numerical data needed. Three responded and provided the requested data. Half of the studies reported moderate to vigorous physical activity (MVPA) outcome measures, which were used for the effect size calculations when available. If the study presented objective measurement of PA, those were used for the
<table>
<thead>
<tr>
<th>Study</th>
<th>Country</th>
<th>Study design</th>
<th>Intervention and control groups</th>
<th>Number of participants</th>
<th>Total attrition</th>
<th>Sex of participants</th>
<th>Mean age of participants (SD)</th>
<th>PA outcomes</th>
<th>SB outcomes</th>
<th>Assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bayne-Smith et al.</td>
<td>US</td>
<td>CRT</td>
<td>1 IV 1 CON</td>
<td>442</td>
<td>Not reported</td>
<td>Only female</td>
<td>-</td>
<td>Self-reported PA no. of sessions per week</td>
<td>Not targeted or reported</td>
<td>Baseline, Post-IV</td>
</tr>
<tr>
<td>Gomes de Barros et al.</td>
<td>Brazil</td>
<td>CRT</td>
<td>1 IV 1 CON</td>
<td>2155</td>
<td>54.1%</td>
<td>55.7% female</td>
<td>18.4 (SD = 2.3)</td>
<td>Self-reported no. of days per week accumulating 60 min of MVPA</td>
<td>Not reported</td>
<td>Baseline, Post-IV</td>
</tr>
<tr>
<td>Hill, Abraham, and</td>
<td>UK</td>
<td>CRT</td>
<td>3 IV 1 CON</td>
<td>620</td>
<td>18.9%</td>
<td>51% female</td>
<td>16.97 (SD = 1.4)</td>
<td>Self-reported 'On average over the last three weeks, I have exercised</td>
<td>Not targeted or reported</td>
<td>Baseline, Post-IV</td>
</tr>
<tr>
<td>Wright (2007)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>energetically for at least 30 minutes ____ times per week.'</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lee, Kuo, Fanaw,</td>
<td>Taiwan</td>
<td>CRT</td>
<td>1 IV 1 CON</td>
<td>94</td>
<td>3.2%</td>
<td>Only female</td>
<td>16.2 (SD = 0.4)</td>
<td>Self-reported MVPA min/week</td>
<td>Not reported</td>
<td>Baseline, Post-IV</td>
</tr>
<tr>
<td>Perng, and Juang (2012)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Change in aerobic walking (steps/day)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lubans and Sylva (2006)</td>
<td>UK</td>
<td>RCT</td>
<td>1 IV 1 CON</td>
<td>78</td>
<td>Post-IV 0%</td>
<td>61.5% female</td>
<td>16.7 (SD = 0.5)</td>
<td>Self-reported PA 'In a typical week, how many days do you do 60 min or more</td>
<td>Not reported</td>
<td>Baseline, Post-3-month follow-up</td>
</tr>
<tr>
<td>Maunielio et al. (2010)</td>
<td>US</td>
<td>CRT</td>
<td>1 IV 1 CON</td>
<td>1800</td>
<td>Post-IV 2.6%</td>
<td>50.8% female</td>
<td>-</td>
<td>MVPA (30-min blocks/day)</td>
<td>Self-reported limited TV viewing</td>
<td>Baseline, Post-IV, 12-month follow-up</td>
</tr>
<tr>
<td>Neumark-Sztainer et al.</td>
<td>US</td>
<td>CRT</td>
<td>1 IV 1 CON</td>
<td>356</td>
<td>Post-IV 3.1%</td>
<td>Only female</td>
<td>-</td>
<td>Self-reported sedentary activity (30-minute blocks/day)</td>
<td>Self-reported sedentary activity (30-minute blocks/day)</td>
<td>Baseline, Post-IV, 9-month follow-up</td>
</tr>
<tr>
<td>Schofield, Mummery,</td>
<td>Australia</td>
<td>CRT</td>
<td>2 IV 1 CON</td>
<td>85</td>
<td>Post-IV 20%</td>
<td>Only female</td>
<td>15.8 (SD 0.8)</td>
<td>Objectively measured PA (pedometer)</td>
<td>Not targeted or reported</td>
<td>Baseline, Mid-IV, Post-IV</td>
</tr>
<tr>
<td>and Schofield (2005)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4-d step count</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Singhal, Misra, Shah,</td>
<td>India</td>
<td>CRT</td>
<td>1 IV 1 CON</td>
<td>209</td>
<td>Post-IV 3.8%</td>
<td>40.2% female</td>
<td>IV 16.04 (SD 0.41)</td>
<td>Self-reported PA &gt; 4 days in a week</td>
<td>Self-reported watching TV (4–5 h/day)</td>
<td>Baseline, Post-IV</td>
</tr>
<tr>
<td>and Gulati (2010)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>CON 16.0 (SD 0.5)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Slootmaker, Chinapaw,</td>
<td>The</td>
<td>RCT</td>
<td>1 IV 1 CON</td>
<td>87</td>
<td>Post-IV 21.8%</td>
<td>63.2% female</td>
<td>15.1 (SD not reported)</td>
<td>Objectively measured MVPA (accelerometer)</td>
<td>Objectively measured sedentary time</td>
<td>Baseline, Post-IV, 8-month follow-up</td>
</tr>
<tr>
<td>Seidell, van Mechelen,</td>
<td>Netherlands</td>
<td>CRT</td>
<td></td>
<td></td>
<td>Post-IV 9.2%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>and Schuit (2010)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

IV: Short for intervention.
CON: Short for control.
effect size calculations; in other cases self-reported outcomes were addressed. When there was no MVPA outcome reported, we used the closest outcome available (see Table 1). We classified the interventions into effective vs. non-effective based on the same PA outcome data that were used in the effect size calculations. Effective trials were defined as those that reported a significant difference ($p < .05$) between intervention and control groups’ PA at first follow-up post-intervention.

We identified which BCTs were present in the effective trials and which BCTs were present in the non-effective trials. We analysed BCT effectiveness using a method modified from one used in Martin et al.’s (2013) systematic review on childhood obesity prevention and management interventions. Effective BCTs were defined as those that were present in a majority (> 50%) of the effective trials but not at all present or present in only one of the non-effective trials. We also attempted to identify the BCTs unique to interventions that did not change PA or SB. We illustrated BCT effectiveness by calculating ‘effectiveness ratios’ of the relative weight of BCTs appearing in two or more trials. The effectiveness ratio was calculated as the ratio of the number of times a BCT was present in an effective intervention divided by the number of times it was present as a component of all interventions, including the ineffective interventions.

**Results**

The search provided 5482 records (see Figure 1) out of which 13 articles reporting 10 unique intervention studies were identified for inclusion in the review (Bayne-Smith et al., 2004; Gomes de Barros et al., 2009; Hill et al., 2007; Lee et al., 2012; Lubans & Sylva, 2006; Mauriello et al., 2006; Mauriello et al., 2010; Nahas et al., 2009; Neumark-Sztainer et al., 2010; Neumark-Sztainer, Story, Hannan, & Rex, 2003; Schofield et al., 2005; Singhal et al., 2010; Slootmaker et al., 2010).

**Study characteristics and participants**

Two of the included studies were individually randomised controlled trials (Lubans & Sylva, 2006; Slootmaker et al., 2010) and eight were cluster randomised trials (Bayne-Smith et al., 2004; Gomes de Barros et al., 2009; Hill et al., 2007; Lee et al., 2012; Mauriello et al., 2010; Neumark-Sztainer et al., 2010; Schofield et al., 2005; Singhal et al., 2010). The average sample size was 593 (SD = 757), ranging from 78 to 2155 older adolescents. The non-weighed mean attrition percentage (where reported) at first follow-up post intervention was 16.2% (SD = 16.8%) ranging from 0% to 54.1%. Four interventions were only aimed at female adolescents (Bayne-Smith et al., 2004; Lee et al., 2012; Neumark-Sztainer et al., 2010; Schofield et al., 2005) and six were aimed at both sexes (Gomes de Barros et al., 2009; Hill et al., 2007; Mauriello et al., 2010; Singhal et al., 2010; Slootmaker et al., 2010). (For more detail on study characteristics, see Table 1.)

**Intervention and comparison arms**

The 10 studies compared a total of 13 different intervention groups against 10 control groups. Eight studies compared one intervention group against one control group (Bayne-Smith et al., 2004; Gomes de Barros et al., 2009; Lee et al., 2012; Lubans & Sylva,
These interventions included a web-based intervention offering personalised feedback on readiness to change (Mauriello et al., 2010), a web-based intervention coupled with the use of accelerometers (Slootmaker et al., 2010), an intervention using pedometers to encourage walking (Lee et al., 2012), an intervention targeting the school environment and offering students PA opportunities (Gomes de Barros et al., 2009), health and exercise programmes (Lubans & Sylva, 2006), a heart health programme focusing on vigorous exercise and other health behaviours such as diet, smoking and stress (Bayne-Smith et al., 2004), and an all-girls PE course targeting self-empowerment, nutrition and including one-on-one motivational interviewing sessions (Neumark-Sztainer et al., 2010) (for further information see Table 2). One study compared two intervention groups against a control group (Schofield et al., 2005). In this study both intervention groups were offered a 12-week PA self-monitoring and educative programme. One intervention group was provided with pedometers and set daily step count targets whereas the other group set daily time-based goals for PA involvement (Schofield et al., 2005). A final study compared three types of theory-based persuasive leaflets to increase PA (leaflet alone, leaflet plus a motivational quiz, and leaflet plus implementation intention prompt) against a no-leaflet control group (Hill et al., 2007). Control group treatment differed widely in the 10 trials. Whereas some control groups received a cognitive word search task (Hill et al., 2007), or a brief leaflet with general PA recommendations (Slootmaker et al., 2010), others received the school’s standard PE programme (Gomes de Barros et al., 2009). Neumark-Sztainer et al. (2010) described also providing the control group with an all-girls PE class. Overall, control group interventions were insufficiently described in the studies to enable coding of BCTs.

**Risk of bias**

All studies except for one (Slootmaker et al., 2010) were judged to be at high risk of bias in at least one domain (see Table 3). However, the study by Slootmaker et al. (2010) also had domains of unclear risk of bias. Sequence generation was inadequately described in eight out of 10 study reports. None of the studies reported proper blinding of participants to study group allocation.

**Outcomes**

Five studies targeted only PA or PA and SB outcomes (Hill et al., 2007; Lee et al., 2012; Lubans & Sylva, 2006; Schofield et al., 2005; Slootmaker et al., 2010). The five other studies targeted other outcomes as well, such as heart health knowledge and dietary behaviours (Bayne-Smith et al., 2004; Gomes de Barros et al., 2009; Mauriello et al., 2010; Neumark-Sztainer et al., 2010; Singhal et al., 2010). Five studies assessed outcome measures only at baseline and at the end of the intervention period (Bayne-Smith et al., 2004; Gomes de Barros et al., 2009; Lee et al., 2012; Schofield et al., 2005; Singhal et al., 2010). Hill et al. (2007) assessed outcomes three weeks post-intervention. Three studies assessed outcomes both post-intervention and a few months post-intervention (Lubans & Sylva, 2006; Neumark-Sztainer et al., 2010; Slootmaker et al., 2010). Mauriello et al.
(2010) reported post-intervention measurements and further follow-ups at 6 and 12 months post-intervention.

All 10 studies reported PA outcomes and four studies reported SB outcomes as well (Mauriello et al., 2010; Neumark-Sztainer et al., 2010; Singhal et al., 2010; Slootmaker et al., 2010). Three other studies mentioned the assessment of SB-related outcomes, but did not report any data (Gomes de Barros et al., 2009; Lee et al., 2012; Lubans & Sylva, 2006). Out of the 10 studies only three used objective PA measurement: two used pedometers (Lee et al., 2012; Schofield et al., 2005) and one used accelerometers (Slootmaker et al., 2010). Seven out of 10 studies relied exclusively on self-report methods to measure PA. Only one out of four studies that measured SB used accelerometry (Slootmaker et al., 2010). The remaining studies used proxy measures of TV viewing (Mauriello et al., 2010); TV viewing, playing board games and attending tuition classes (Singhal et al., 2010); and blocks of sedentary activity using the Self-reported Total Physical Activity (3-DPAR) questionnaire (Neumark-Sztainer et al., 2010).

### Intervention effects on physical activity and sedentary behaviours

Seven studies showed significant differences in PA between intervention and control groups post-intervention (Gomes de Barros et al., 2009; Hill et al., 2007; Lee et al., 2012; Lubans & Sylva, 2006; Mauriello et al., 2010; Schofield et al., 2005; Slootmaker et al., 2010). The effect sizes (Cohen’s d) post-intervention ranged from small to medium (0.132–0.659). Three studies did not find significant between-group differences (Bayne-Smith et al., 2004; Neumark-Sztainer et al., 2010; Singhal et al., 2010). None of the four studies (Lubans & Sylva, 2006; Mauriello et al., 2010; Neumark-Sztainer et al., 2010; Slootmaker et al., 2010) with follow-ups beyond one month post-intervention reported significant differences between groups, although three of the four studies (Lubans & Sylva, 2006; Mauriello et al., 2010; Slootmaker et al., 2010) observed significant effects post-intervention.

Out of the four studies that measured SB, one found significant differences between intervention and control groups in self-reported sedentary activity over a nine-month post-baseline follow-up (Neumark-Sztainer et al., 2010). Slootmaker et al. (2010) reported a significant difference between intervention and control groups in objectively measured sedentary time at eight months follow-up. The other two studies showed no significant differences between groups in SB measured by time spent viewing television (Mauriello et al., 2010; Singhal et al., 2010). Because of the lack of studies that focused solely on SB, analysis of intervention characteristics and BCTs related to effectiveness in this review will focus on PA only.

### Intervention characteristics

Table 2 describes intervention characteristics. Five of seven interventions that increased PA targeted PA or PA and SB only (Hill et al., 2007; Lee et al., 2012; Lubans & Sylva, 2006; Schofield et al., 2005; Slootmaker et al., 2010), whereas two targeted multiple behaviours related to energy balance, including dietary behaviour (Gomes de Barros et al., 2009; Mauriello et al., 2010). All three interventions that did not increase PA targeted multiple behaviours: two targeted obesity-related behaviours (Neumark-Sztainer et al., 2010; Mauriello et al., 2010). One study targeted PA and SB (Mauriello et al., 2010), and one study targeted PA (Lee et al., 2012). The remaining two studies focused on PA and SB (Hill et al., 2007; Singhal et al., 2010).

### Table 2

<table>
<thead>
<tr>
<th>Study</th>
<th>Intervention Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hill et al., 2007</td>
<td>Increased PA and SB&lt;br&gt;Targeted PA and SB&lt;br&gt;Used self-report measures</td>
</tr>
<tr>
<td>Lee et al., 2012</td>
<td>Increased PA and SB&lt;br&gt;Targeted PA and SB&lt;br&gt;Used self-report measures</td>
</tr>
<tr>
<td>Lubans &amp; Sylva, 2006</td>
<td>Increased PA and SB&lt;br&gt;Targeted PA and SB&lt;br&gt;Used self-report measures</td>
</tr>
<tr>
<td>Schofield et al., 2005</td>
<td>Increased PA and SB&lt;br&gt;Targeted PA and SB&lt;br&gt;Used self-report measures</td>
</tr>
<tr>
<td>Slootmaker et al., 2010</td>
<td>Increased PA and SB&lt;br&gt;Targeted PA and SB&lt;br&gt;Used self-report measures</td>
</tr>
<tr>
<td>Neumark-Sztainer et al., 2010</td>
<td>Increased PA and SB&lt;br&gt;Targeted PA and SB&lt;br&gt;Used self-report measures</td>
</tr>
<tr>
<td>Mauriello et al., 2010</td>
<td>Increased PA and SB&lt;br&gt;Targeted PA and SB&lt;br&gt;Used self-report measures</td>
</tr>
<tr>
<td>Lee et al., 2012</td>
<td>Increased PA and SB&lt;br&gt;Targeted PA and SB&lt;br&gt;Used self-report measures</td>
</tr>
<tr>
<td>Gomes de Barros et al., 2009</td>
<td>Increased PA and SB&lt;br&gt;Targeted PA and SB&lt;br&gt;Used self-report measures</td>
</tr>
<tr>
<td>Singhal et al., 2010</td>
<td>Increased PA and SB&lt;br&gt;Targeted PA and SB&lt;br&gt;Used self-report measures</td>
</tr>
<tr>
<td>Mauriello et al., 2010</td>
<td>Increased PA and SB&lt;br&gt;Targeted PA and SB&lt;br&gt;Used self-report measures</td>
</tr>
<tr>
<td>Slootmaker et al., 2010</td>
<td>Increased PA and SB&lt;br&gt;Targeted PA and SB&lt;br&gt;Used self-report measures</td>
</tr>
</tbody>
</table>

Downloaded by [Tampere University] at 22:47 08 June 2016
### Table 2. Intervention characteristics.

<table>
<thead>
<tr>
<th>Study</th>
<th>Mean values (SD) on which Cohen’s $d$ was calculated</th>
<th>Risk of bias (number of items)</th>
<th>Intervention targets</th>
<th>Intervention length</th>
<th>Intervention delivery mode</th>
<th>Intervention provider</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lubans and Sylva (2006)</td>
<td>MVPA min/week IV: 250.5 (113.4) CON: 172.5 (123.3)</td>
<td>High risk: 3 Unclear risk: 2  Low risk: 1 N/A: 2</td>
<td>Physical activity</td>
<td>~ 10 weeks</td>
<td>Face-to-face groups, self-led exercise</td>
<td>1 member of the research team</td>
</tr>
<tr>
<td>Hill et al. (2007)</td>
<td>LII:0.586* L: 0.484 LQ:0.384 Exercise times/wk: LII: 3.17 (2.0) L: 2.96 (1.9) LQ: 3.46 (2.3) CON: 2.52 (1.80)</td>
<td>High risk: 2 Unclear risk: 0  Low risk: 6</td>
<td>Physical activity</td>
<td>1 session</td>
<td>Leaflets, written material</td>
<td>Research assistant</td>
</tr>
<tr>
<td>Schofield et al. (2005)</td>
<td>S: 0.502** T: 0.241 Aerobic steps/day IV: 836 (832.03) CON: 515 (554.19)</td>
<td>High risk: 3 Unclear risk: 2  Low risk: 1</td>
<td>Physical activity</td>
<td>12 weeks</td>
<td>Face-to-face small groups, written materials e.g. log- &amp; textbooks, pedometers</td>
<td>Research staff (principal researcher or assistant)</td>
</tr>
<tr>
<td>Lee et al. (2012)</td>
<td>0.454 Aerobic steps/day IV: 836 (832.03) CON: 515 (554.19)</td>
<td>High risk: 1 Unclear risk: 3  Low risk: 4</td>
<td>Physical activity</td>
<td>12 weeks</td>
<td>Face-to-face group and individual discussions, pedometers</td>
<td>1 researcher with public health nursing background</td>
</tr>
<tr>
<td>Gomes de Barros et al. (2009)</td>
<td>0.333 No. of days/week accumulating 60 min of MVPA IV: 3.3 (2.1) CON: 2.6 (2.1)</td>
<td>High risk: 1 Unclear risk: 3  Low risk: 4</td>
<td>Multiple behaviours</td>
<td>9 months</td>
<td>Face-to-face groups, written materials</td>
<td>Trained school staff</td>
</tr>
<tr>
<td>Slootmaker et al. (2010)</td>
<td>0.174 –</td>
<td>High risk: 0 Unclear risk: 2  Low risk: 4 N/A: 0</td>
<td>Physical activity</td>
<td>12 weeks</td>
<td>Individualised Internet-based programme, accelerometers</td>
<td>Intervention research group</td>
</tr>
<tr>
<td>Mauriello et al. (2010)</td>
<td>0.132 –</td>
<td>High risk: 5 Unclear risk: 1  Low risk: 2 N/A: 0</td>
<td>Multiple behaviours</td>
<td>~ 8 weeks</td>
<td>Interactive computer program</td>
<td>Research assistant</td>
</tr>
<tr>
<td>Neumark-Sztainer et al. (2010)</td>
<td>Not sign.</td>
<td>High risk: 2 Unclear risk: 4  Low risk: 2</td>
<td>Multiple behaviours</td>
<td>~ 16 weeks</td>
<td>Face-to-face groups &amp; individual meetings, textbooks and postcards</td>
<td>Trained PE teachers, Community guest instructors, Research staff</td>
</tr>
<tr>
<td>Singhal et al. (2010)</td>
<td>Not sign.</td>
<td>High risk: 4 Unclear risk: 2  Low risk: 2</td>
<td>Multiple behaviours</td>
<td>24 weeks</td>
<td>Face-to-face groups &amp; individual meetings</td>
<td>Trained nutritionist, unclear</td>
</tr>
<tr>
<td>Bayne-Smith et al. (2004)</td>
<td>Not sign.</td>
<td>High risk: 5 Unclear risk: 3  Low risk: 0</td>
<td>Multiple behaviours</td>
<td>12 weeks</td>
<td>Face-to-face groups, written materials, homework assignments</td>
<td>Trained PE teachers</td>
</tr>
</tbody>
</table>

IV: Short for Intervention Group.
CON: Short for Control Group.
Table 3. Risk of bias summary.

<table>
<thead>
<tr>
<th>Risk of bias summary</th>
<th>Adequate sequence generation</th>
<th>Allocation concealment</th>
<th>Blinding (Subjective outcomes)</th>
<th>Incomplete outcome data addressed</th>
<th>Free of selective outcome reporting</th>
<th>Free of other sources of bias</th>
<th>Free of cluster recruitment bias</th>
<th>Free of baseline imbalance bias</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lubans and Sylva (2006)</td>
<td>unclear</td>
<td>unclear</td>
<td>high</td>
<td>low</td>
<td>high</td>
<td>high</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Hill et al. (2007)</td>
<td>low</td>
<td>low</td>
<td>high</td>
<td>low</td>
<td>low</td>
<td>high</td>
<td>low</td>
<td>low</td>
</tr>
<tr>
<td>Schofield et al. (2005)</td>
<td>unclear</td>
<td>unclear</td>
<td>high</td>
<td>high</td>
<td>low</td>
<td>high</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Lee et al. (2012)</td>
<td>unclear</td>
<td>unclear</td>
<td>high</td>
<td>low</td>
<td>low</td>
<td>unclear</td>
<td>low</td>
<td>low</td>
</tr>
<tr>
<td>Gomes de Barros et al. (2009)</td>
<td>unclear</td>
<td>unclear</td>
<td>high</td>
<td>low</td>
<td>low</td>
<td>uncertain</td>
<td>high</td>
<td>low</td>
</tr>
<tr>
<td>Slootmaker et al. (2010)</td>
<td>low</td>
<td>low</td>
<td>unclear</td>
<td>low</td>
<td>low</td>
<td>uncertain</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Mauriello et al. (2010)</td>
<td>unclear</td>
<td>high</td>
<td>high</td>
<td>low</td>
<td>low</td>
<td>high</td>
<td>high</td>
<td>high</td>
</tr>
<tr>
<td>Neumark-Sztainer et al. (2010)</td>
<td>unclear</td>
<td>unclear</td>
<td>high</td>
<td>unclear</td>
<td>low</td>
<td>high</td>
<td>unclear</td>
<td>low</td>
</tr>
<tr>
<td>Singhal et al. (2010)</td>
<td>unclear</td>
<td>unclear</td>
<td>high</td>
<td>low</td>
<td>low</td>
<td>high</td>
<td>high</td>
<td>high</td>
</tr>
<tr>
<td>Bayne-Smith et al. (2004)</td>
<td>unclear</td>
<td>high</td>
<td>high</td>
<td>unclear</td>
<td>unclear</td>
<td>high</td>
<td>high</td>
<td>high</td>
</tr>
</tbody>
</table>
Singhal et al., 2010) and one targeted heart health knowledge and behaviours (Bayne-Smith et al., 2004). To summarise, it appears that effective interventions focus on PA or PA and SB only.

**Length of intervention**

Intervention length ranged from one intervention session (Hill et al., 2007) to nine months (Gomes de Barros et al., 2009). The median length of intervention was 12 weeks. Intervention length appeared unrelated to intervention effectiveness.

**Mode of intervention delivery**

Two interventions were delivered via an interactive, individualised computer program (Mauriello et al., 2010; Slootmaker et al., 2010), and one solely via leaflets (Hill et al., 2007). The remaining seven interventions were delivered face-to-face in groups (Bayne-Smith et al., 2004; Gomes de Barros et al., 2009; Lee et al., 2012; Lubans & Sylva, 2006; Neumark-Sztainer et al., 2010; Schofield et al., 2005; Singhal et al., 2010), some including individual meetings with adolescents (Lee et al., 2012; Neumark-Sztainer et al., 2010; Singhal et al., 2010). All of the interventions delivered face-to-face also included written materials, such as textbooks, logbooks, postcards and homework assignments. Two interventions used pedometers (Lee et al., 2012; Schofield et al., 2005) and one study used accelerometers (Slootmaker et al., 2010) as an intervention strategy. There appeared to be no systematic differences in delivery mode between effective versus non-effective interventions.

**Intervention provider**

Two out of 10 interventions were delivered by trained school staff and PE teachers (Bayne-Smith et al., 2004; Gomes de Barros et al., 2009) and seven by researchers (Hill et al., 2007; Lee et al., 2012; Lubans & Sylva, 2006; Mauriello et al., 2010; Schofield et al., 2005; Singhal et al., 2010; Slootmaker et al., 2010). One study utilised trained PE teachers, community guests and intervention staff for programme delivery (Neumark-Sztainer et al., 2010). Six of the seven effective interventions were delivered by researchers (Hill et al., 2007; Lee et al., 2012; Lubans & Sylva, 2006; Mauriello et al., 2010; Schofield et al., 2005; Slootmaker et al., 2010). Of the three ineffective interventions, one was delivered by school staff only (Bayne-Smith et al., 2004), the second by trained school staff, community guest instructors and research staff (Neumark-Sztainer et al., 2010) and the third was delivered by a trained nutritionist, but details were lacking about potential other providers (Singhal et al., 2010). To summarise, most of the effective interventions were delivered by researchers rather than school staff.

**Intervention content: BCTs**

Table 4 presents BCTs identified in more than one of the included studies. Full details on all of the identified BCTs are available from the authors. Both coders agreed that 57 out of the 93 BCTs in the BCTT v1 were not present in any intervention description. For the remaining 36 BCTs inter-rater reliability was good for 28 BCTs (percentage agreement ranging from 75% to 100%) and sub-optimal (< 75%) for the remaining eight, due to their infrequent inclusion in the descriptions. The coders were unable to conduct proper analysis of the BCTs used in control groups due to insufficient reporting in the research articles.
<table>
<thead>
<tr>
<th>Table 4. Behaviour Change Techniques present in interventions.</th>
</tr>
</thead>
<tbody>
<tr>
<td>--------------------</td>
</tr>
<tr>
<td><strong>Instruction on how to perform a behaviour</strong></td>
</tr>
<tr>
<td><strong>Goal setting (behaviour)</strong></td>
</tr>
<tr>
<td><strong>Demonstration of the behaviour</strong></td>
</tr>
<tr>
<td><strong>Action planning</strong></td>
</tr>
<tr>
<td><strong>Behavioural practice/rehearsal</strong></td>
</tr>
<tr>
<td><strong>Social support (unspecified)</strong></td>
</tr>
<tr>
<td><strong>Information about social and environmental consequences</strong></td>
</tr>
<tr>
<td><strong>Graded tasks</strong></td>
</tr>
<tr>
<td><strong>Feedback on behaviour</strong></td>
</tr>
<tr>
<td><strong>Self-monitoring of behaviour</strong></td>
</tr>
<tr>
<td><strong>Problem solving</strong></td>
</tr>
<tr>
<td><strong>Social comparison</strong></td>
</tr>
<tr>
<td><strong>Review behaviour goals</strong></td>
</tr>
<tr>
<td><strong>Discrepancy between current behaviour and goal</strong></td>
</tr>
<tr>
<td><strong>Restructuring the physical environment</strong></td>
</tr>
</tbody>
</table>

**BCT frequency**

1. Cohen's $d$
2. Lubans and Sylva (2006)
3. Hill et al. (2007)
4. Schofield et al. (2005)
5. Lee et al. (2012)
6. Gomes de Barros et al. (2009)
7. Slootmaker et al. (2010)
8. Mauriello et al. (2010)
An average of 10.5 (range = 5–20) BCTs were included in effective interventions, and an average of four BCTs (range = 3–6) in ineffective interventions. BCTs unique to effective interventions were Information about social and environmental consequences, Graded tasks, Self-monitoring of behaviour, Feedback on behaviour, Problem solving, Goal setting (behaviour), Action planning and Social support (unspecified).

Behavioural practice was present in all three ineffective trials but only in two out of seven effective interventions. Instruction on how to perform a behaviour and Demonstration of the behaviour were frequently present in both effective and non-effective interventions. Figure 2 describes the ratio of effectiveness for BCTs identified two or more times in studies included in this review.

Discussion

To our knowledge, this is the first systematic review to investigate school-based interventions targeting PA and SB among 15–19-year-old adolescents, characterising intervention content using the BCT Taxonomy v1, and linking intervention characteristics to intervention effectiveness. In total, 10 trials met the inclusion criteria and seven of these increased PA in the short term, but effects were not sustained at longer follow-ups. The trials with significant effects on PA focused on PA or PA and SB only and tended to use researchers as providers. Effectiveness was unrelated to length of intervention or mode of delivery. Effective interventions, and especially those with an effect size ranging from medium to
large, reported using more BCTs than the non-effective interventions, and specifically eight BCTs: *Information about social and environmental consequences, Graded tasks, Self-monitoring of behaviour, Feedback on behaviour, Problem solving, Goal setting (behaviour), Action planning* and *Social support*. Out of the four studies that measured SB, only two had significant effects on the behaviour. No studies were found that focused solely on SB. Due to this lack of evidence, it is difficult to draw conclusions on what intervention elements or BCTs were related to effectiveness in regards to SB.

Our review suggests that interventions targeting multiple health behaviours are less effective in promoting PA among older adolescents than ones focusing solely on PA or PA and SB, which is also supported by earlier evidence (Crutzen et al., 2010). Previous reviews have suggested that when appropriately trained, a wide range of providers can deliver effective health behaviour interventions (Greaves et al., 2011; Peters, Kok, Ten Dam, Buijs, & Paulussen, 2009). On the contrary, we found that interventions delivered by research staff were more effective than those delivered by school staff or other providers. This finding resembles the basic difference between efficacy and effectiveness trials, and may be more a matter of fidelity of delivery – or training – rather than a matter of the person providing the intervention. While other reviews have shown that longer duration of school-based PA interventions was associated with effectiveness (Dobbins et al., 2009; Safron et al., 2011), we did not find a relationship between intervention length and short-term effectiveness. It is possible that the relationship between intervention length and PA outcomes was influenced by other intervention characteristics not analysed in our review such as intervention intensity: short but very intensive interventions may be more effective than longer interventions with less contact. Finally, in line with previous evidence (Greaves et al., 2011), there was no clear association between mode of intervention delivery and effectiveness.

In line with some (Avery, Flynn, van Wersch, Sniehotta, & Trenell, 2012; Hendrie et al., 2012) and in contrast to other reviews (Abraham & Graham Rowe, 2009; Dombrowski et al., 2013), we found that effective PA interventions included more BCTs than non-effective interventions. This may be either a matter of non-effective interventions actually using fewer techniques to change behaviour, or of not reporting what was done, indicating less precision in both planning and reporting the intervention. The specific BCTs identified to be related with intervention effectiveness resonate with the findings of previous reviews with adult populations (Avery et al., 2012; Michie et al., 2009; Williams & French, 2011), where BCTs related to self-regulation have been shown to be effective in changing PA. Based on our review, this applies also among older adolescents: *Goal setting* and *Action planning* were techniques emphasised in effective trials, and complementing BCTs such as *Self-monitoring of behaviour, Feedback on behaviour* and *Problem solving* were unique to effective trials.

In their review on obesity prevention interventions for children, Hendrie et al. (2012) found that providing information on behaviour–health links was related to intervention effectiveness. Quite similarly, we found *Information about social and environmental consequences* to be a technique uniquely present in the effective trials. What should be noted is that this BCT is coded whenever the information provided is unspecified in the intervention description (Michie et al., 2013). Therefore it is possible that the information provided in the interventions was not actually about social and environmental consequences, but rather about health consequences or emotional consequences, or a combination of these.
Interestingly, *Behavioural practice/rehearsal* was a technique present mainly in non-effective interventions. In a previous review of PA interventions among obese adults (Olander et al., 2013), quite to the contrary, *Prompting practice* was related to largest effects for PA. It is possible that prompting behavioural practice outside of intervention sessions is related to effectiveness whereas practice within the sessions is a typical element of non-effective interventions. This is supported by our finding that setting *Graded tasks* was a BCT present uniquely in effective interventions.

We found little evidence of intervention techniques focusing on the environment level of behaviour change, such as *Restructuring the physical environment*. Neither have previous reviews on PA interventions (see, e.g., Greaves et al., 2011; Michie et al., 2009) shown these techniques to be related to effectiveness. *Restructuring the physical environment* was present in only three interventions, two of which were effective and one non-effective. The lack of environment-level techniques may be a matter of financial resources. Initiating major changes in the school’s physical environment without previous evidence of it actually being effective in supporting and increasing students’ PA might be too risky and costly. Several distinct environmental changes, such as providing opportunities for PA by direct monetary investments in schools’ PA equipment, offering students free opportunities for structured PA and organising collective supervised PA events, were heavily emphasised in the intervention by Gomes de Barros et al. (2009). However, in the BCT Taxonomy v1 all of these different actions are coded under one technique: *Restructuring the physical environment*. We recommend that this technique is unpacked in more detail in future developments of the taxonomy, so that studies can investigate the relative effectiveness of the above approaches.

Out of the four studies that measured sedentary behaviour, only two significantly decreased SB. Based on only two studies it is difficult to draw reliable conclusions on what elements were related to effectiveness. Notably, none of the trials measuring SB treated it as a separate behaviour, but rather as an indicator of insufficient PA levels. SB was measured objectively in only one of the studies (Slootmaker et al., 2010). The others relied on self-reported SB often indicated by time spent on the computer, viewing television or playing board games. The studies provided no descriptions of intervention strategies aimed specifically at changing SB, which made it impossible to identify the BCTs effective in reducing SB, and make a distinction with PA. The BCTs and alterations in the school context that may need to occur in order to decrease adolescents’ SB may be quite different from the strategies needed to enhance PA. Ultimately, an intervention focusing on PA may not affect SB (e.g., if the focus of the intervention is on increasing the intensity of the PA). Conversely, an SB intervention may not affect health-enhancing MVPA if adolescents are encouraged to substitute sitting with standing or light-intensity PA. It is evident that more studies are needed that target intervention strategies to specifically decrease SB.

Furthermore, in order to reliably identify the BCTs associated with PA and SB intervention effectiveness, the BCTs used in control groups, i.e. standard care, should be described in intervention reports and coded. Especially when the intervention is delivered during the PE class, it is difficult to draw conclusions on effective BCTs without knowing which BCTs are also part of standard care. Control group treatment differed widely in the 10 trials and was altogether insufficiently described in the studies. Neumark-Sztainer et al. (2010) described providing the control group with an all-girls PE class with mostly inactive participants. It is possible that this alone is quite a powerful intervention, deviating from
the standard mixed-gender PE, and may explain the lack of statistically significant differences in PA between intervention and control groups post-intervention. We recommend that future studies report BCTs used in both intervention and control groups.

It is notable that four out of 10 interventions targeted females exclusively. Girls are indeed more in need of PA-promoting interventions, as they engage on average in less MVPA than boys (Sallis, Prochaska, & Taylor, 2000). However, boys generally spend more time on screen-based behaviours. Previous studies have shown that school-based PA interventions work better for girls than boys (Yildirim et al., 2011), which may be in part due to girls’ low baseline MVPA levels.

The limitations of our review are influenced by the number and quality of the studies included. In general, the methodological quality of the included studies was weak, with relatively small sample sizes and a high risk of bias. Only three studies reported objective measurement of PA (accelerometry, pedometers) and one study (Slootmaker et al., 2010) objectively assessed SB, posing a major reliability problem. Out of the seven studies that we classified as effective in increasing PA, four relied solely on self-report measures. Since self-report can be influenced by recall bias and social desirability bias, these findings need to be confirmed by objective measures. Another major limitation in this area, as noted in earlier systematic reviews as well (e.g., van Sluijs et al., 2007), is the lack of long-term follow-ups and, consequently, a lack of evidence regarding long-term effectiveness. Thirdly, delivering interventions only to volunteer participants expressing heightened interest in the programme (Lubans & Sylva, 2006; Mauriello et al., 2010) versus, for example, recruiting all students from a given classroom possesses a threat of positively biased results. It is quite evident that individuals who seek to participate in a PA intervention are more motivated to increase their level of PA than their age cohort in general. Also, because self-reported PA was the main outcome measure in most studies, it is possible that failure to blind participants to allocation might have resulted in differential effects of social desirability bias between intervention and control group participants. There was an apparent lack of quality in reporting the studies. Some of the studies were so superficially described that it was difficult to draw conclusions on intervention elements related to effectiveness; from some it was difficult to even tell who delivered the intervention. However, the inter-rater reliability of the coders was adequate. Finally, due to the small number of included studies and the heterogeneity of the outcome measures, we were not able to conduct a meta-analysis. Hence, it should be taken into account that the conclusions on associations and relationships between intervention components and intervention effectiveness presented in this review are based on a small number of studies and descriptive rather than quantitative analysis. The scarcity of reported studies in this age group and the fact that only three non-effective studies were found in the literature search also raises a serious question of possible risk of publication bias in this research field. Due to these factors, the conclusions drawn from this review regarding BCT effectiveness are tentative.

This study has several implications for future research and practice. We recommend that future intervention studies use objective measures of PA and SB alongside self-report measures and use large enough sample sizes to allow for multi-level analyses to account for clustering within classes and schools. Furthermore, future trials should include long-term follow-ups and more precise descriptions of intervention content in terms of BCTs.

The advantages of school-based PA and SB interventions include a captive audience and wide reach of the target group. As there is a marked drop in PA levels during
adolescence, and a high level of SB, this age group is an especially important target for public health efforts. Our review may help efforts to develop more effective interventions in this under-researched age group, having identified factors associated with effectiveness and crucial gaps in research that merit attention in future studies. First, for future intervention developers, we would point out the major lack of school-based intervention studies to decrease SB in this age group. Future studies ought to acknowledge that PA and SB are different target behaviours that require different intervention strategies. In our review, we identified promising BCTs that seem to be effective at least in the short term in school-based interventions to promote PA among older adolescents. It should be noted that interventions should not be deemed effective based on statistical significance only, but one should also examine whether the changes have public health or clinical significance. Few of the original studies made an explicit judgement, or assessed cost-effectiveness. Since our review found that interventions delivered via research staff were more effective than those delivered by school staff or other providers, we suggest that in future research careful process evaluation (Moore et al., 2015) is conducted to examine acceptability and fidelity of the intervention and to understand whether public health interventions can be rolled out successfully without researcher delivery. In addition, intervention developers should carefully select BCTs and other elements that would increase maintenance of PA over the long term, as so far no study has been able to demonstrate that their school-based intervention have long-lasting effects on older adolescents’ PA. Since our review shows a lack of evidence in this target audience and behaviour, we suggest that developers look into other areas of behavioural science research where BCTs aimed at maintaining behaviour are being tested.

In conclusion, this review found that there is limited evidence on how to best promote PA and reduce SB among older adolescents in school-based interventions in the long term. The method we used to analyse the content of interventions was helpful in identifying effective elements, which will benefit future intervention development work, ultimately improving adolescent activity behaviours.

Disclosure statement
No potential conflict of interest was reported by the authors.

Funding
The first author S-T. Hynynen was supported by the Ministry of Education and Culture [grant number 34/626/2012], and the Ministry of Social Affairs and Health [grant number: 201310238]; Finland. F. F. Sniehotta is funded by Fuse, the UK Clinical Research Collaboration Centre of Excellence for Translational Research in Public Health. Funding for Fuse from the British Heart Foundation, Cancer Research UK, Economic and Social Research Council, Medical Research Council, and the National Institute for Health Research, under the auspices of the UK Clinical Research Collaboration.

Supplemental data
Supplemental data for this article can be accessed 10.1080/1750984X.2015.1081706.
Note
1. Prompting practice is a BCT in the CALO-RE taxonomy; a refined taxonomy of behaviour change techniques to help people change their physical activity and healthy eating behaviour (see Michie et al. 2011)

References


activity targets and behaviour change techniques are associated with intervention effectiveness? *Obesity Reviews*, 12(2), 114–130.


Michie, S., Richardson, M., Johnston, M., Abraham, C., Francis, J., Hardeman, W., ... Wood, C. E. (2013). The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: Building


Williams, S. L., & French, D. P. (2011). What are the most effective intervention techniques for changing physical activity self-efficacy and physical activity behaviour—and are they the same? *Health Education Research, 26*(2), 308–322.
