Health Education
& Behavior

Physically active lessons improve lesson activity and on-task behaviour: A cluster-randomised controlled trial of the 'Virtual Traveller' intervention

Journal:	*Health Education & Behavior*
Manuscript ID	HEB-17-0246.R1
Manuscript Type:	Original Manuscript
Keywords:	Physically Active Lessons, Physical Activity / Exercise, On-task behaviour, Student engagement, Intervention, School-Based Health Care

Abstract:

Background: Physically active lessons have not often been assessed with randomised controlled trials.
Aims: Evaluate the effects of the 'Virtual Traveller' intervention delivered using classroom interactive whiteboards on physical activity, on-task behaviour and student engagement.
Methods: Participants were 219 children aged 8-9 years from ten schools in Greater London, assessed in a cluster-randomised controlled trial between March 2015 and May 2016. For six weeks, intervention children received 10-minute 'Virtual Traveller' sessions three times a week during maths and English lessons (VT group: n=113). Children in control schools received regular teaching (COM group: n=106). Outcomes were school-day, weekend-day and lesson-time sedentary behaviour (SB), light (LPA) and moderate-to-vigorous physical activity (MVPA), on-task behaviour and student engagement, assessed at baseline (T0), two- (T1) and four weeks (T2) during the Virtual Traveller intervention and one week (T3) and three months (T4) post-intervention using multilevel modelling.
Results: VT pupils engaged in significantly more school-day MVPA at T1 only, with no other significant differences between groups in overall school-day or weekend-day activity. VT pupils engaged in significantly less SB and more MVPA during lesson time than COM pupils. More on-task behaviour was shown in VT pupils than COM pupils but there was no difference in student engagement.
Discussion: Virtual Traveller reduced sedentary behaviour and increased physical activity during lesson time but not across overall school or weekend-days. It improved on-task behaviour but had no effect on student engagement.
Conclusion: Physical activity can be integrated into teaching using interactive whiteboards with no detriment to educational outcomes.

Physically active lessons improve lesson activity and on-task behaviour:

A cluster-randomised controlled trial of the 'Virtual Traveller' intervention

Abstract

Background: Physically active lessons have not often been assessed with randomised controlled trials.

Aims: Evaluate the effects of the 'Virtual Traveller' intervention delivered using classroom interactive whiteboards on physical activity, on-task behaviour and student engagement.

Methods: Participants were 219 children aged 8-9 years from ten schools in Greater London, assessed in a cluster-randomised controlled trial between March 2015 and May 2016. For six weeks, intervention children received 10-minute 'Virtual Traveller' sessions three times a week during maths and English lessons (VT group: n=113). Children in control schools received regular teaching (COM group: n=106). Outcomes were school-day, weekend-day and lesson-time sedentary behaviour (SB), light (LPA) and moderate-to-vigorous physical activity (MVPA), on-task behaviour and student engagement, assessed at baseline (T0), two- (T1) and four weeks (T2) during the Virtual Traveller intervention and one week (T3) and three months (T4) post-intervention using multilevel modelling.

Results: VT pupils engaged in significantly more school-day MVPA at T1 only, with no other significant differences between groups in overall school-day or weekend-day activity. VT pupils engaged in significantly less SB and more MVPA during lesson time than COM pupils. More on-task behaviour was shown in VT pupils than COM pupils but there was no difference in student engagement.

Discussion: Virtual Traveller reduced sedentary behaviour and increased physical activity during lesson time but not across overall school or weekend-days. It improved on-task behaviour but had no effect on student engagement.

Conclusion: Physical activity can be integrated into teaching using interactive whiteboards with no detriment to educational outcomes.

Introduction

Physical activity has been shown as beneficial to children's cardiometabolic health (Cesa et al., 2014; Stamatakis et al., 2015), mental health (Biddle & Asare, 2011), cognitive function (Carson et al., 2015) and academic achievement (Efrat, 2011; Fedewa & Ahn, 2011). However the majority of children's time is sedentary, with up to 8.6 hours a day spent in obligatory seated school lessons (LeBlanc et al., 2015). As childhood levels of physical activity (Telama, 2009) and sedentary behaviour (Biddle, Pearson, Ross, & Braithwaite, 2010) have been shown to track into later life, it is vital that interventions are developed to help encourage active lifestyles at an early age (Weiler, Allardyce, Whyte, & Stamatakis, 2013). Various interventions have been developed to add physical activity into the school environment (Dobbins, Husson, DeCorby, & LaRocca, 2013), including during break times (Engelen et al., 2013) and educational sessions (Turner & Chaloupka, 2017). However, teachers typically describe a lack of time as the primary barrier for physical activity provision (Naylor et al., 2015), with such interventions often requiring time to be drawn away from other academic objectives.

To address low activity levels and maintain maximal teaching time; lessons that incorporate physical activity in the teaching of academic content have recently been developed and tested (Mullender-Wijnsma et al., 2016; Norris, Shelton, Dunsmuir, Duke-Williams, & Stamatakis, 2015a). These lesson interventions have typically reported increases to school time physical activity (Mullender-Wijnsma et al., 2015; Riley, Lubans, Holmes, & Morgan, 2016); however follow-up is often limited and the activity measurement used is usually poor (Norris et al., 2015a). Only the 'Physical Activity Across the Curriculum' (PAAC) randomised controlled trial (RCT) has assessed activity levels beyond school time only (Donnelly et al.,

2009; Norris et al., 2015a), finding weekday and weekend accelerometer-assessed activity to be increased at 3-year follow-up (Donnelly et al., 2009). There is hence an unclear evidence base as to whether physically active lessons have effects on activity beyond school time.

Promising educational benefits are evident in initial physically active lesson research (Norris et al., 2015a). The recent 'Fit & Vaardig op School' (Fit and Academically Proficient at School (F&V)) intervention found significant improvements to maths and spelling tests at 2-year follow-up, equating to four months increased learning gains compared to control group (Mullender-Wijnsma et al., 2016). However, wider educational outcomes which influence academic achievement test scores (Finn & Zimmer, 2012; Howie & Pate, 2012) have not been robustly assessed via RCTs. For example, student engagement (behaviour and cognitions in pupils that reflect their interest in learning and school) (Finn & Zimmer, 2012; Fredricks et al., 2011) has not been assessed in relation to physically active lessons. Affective student engagement (emotional connectedness to the school environment) and cognitive student engagement (level of perceived capability and investment towards education) (Fredricks et al., 2011) have been unexplored: meaning that important pupil cognitions towards learning in the context of physically active lessons are still unclear. On-task behaviour during lesson times as a measure of behavioural student engagement (motor and verbal behaviour appropriate to learning situations) (Grieco, Jowers, & Bartholomew, 2009) is not commonly assessed in active lesson RCTs (Mullender-Wijnsma et al., 2015; Norris et al., 2015a). Previous active lesson interventions have mostly not described their behaviour change techniques (BCTs) (Martin & Murtagh, 2015): the 'active ingredients' of intervention content included to encourage a change in behaviour (Michie, Fixsen, Grimshaw, & Eccles,

2009; Michie et al., 2013). Also, physically active lesson research has largely not utilised existing classroom equipment of interactive whiteboards (Norris et al., 2015a; Reference blinded), available in over 70% of UK classrooms (Futuresource Consulting, 2010). This is despite other research showing physical activity to be increased with the provision of other digital technologies, such as Active Video Games (Norris, Hamer, & Stamatakis, 2016; Peng, Crouse, & Lin, 2013).

The aim of this study was to test the effect of the 'Virtual Traveller' intervention on children's physical activity and sedentary behaviour, on-task behaviour and student engagement. The Virtual Traveller (VT) intervention was developed as a series of sessions to incorporate physical activity into primary school maths and English teaching (Reference blinded). It featured a package of pre-prepared Powerpoint sessions delivered by classroom teachers on existing classroom interactive whiteboards. Following recommendations for the development and evaluation of complex health interventions by the Medical Research Council (Medical Research Council, 2013); VT was developed following iterative feasibility work in the form of a pilot study (Reference blinded) and qualitative teacher interviews and pupil focus groups (Reference blinded). It was hypothesised that Virtual Traveller would: 1) increase children's light- (LPA) and moderate-to-vigorous physical activity (MVPA) and reduce sedentary behaviour (SB) during school time, 2) increase LPA and MVPA and reduce SB during lesson time and 3) improve on-task behaviour during lesson time (Reference blinded). This study is reported in accordance with the Consolidation Standards of Reporting Trials (CONSORT) guidelines (Schulz, Altman, & Moher, 2010).

Methods

Design and recruitment

A summary of the study protocol is presented here, with a full version available in the published protocol paper (Reference blinded). The study was a cluster-RCT of a physically active lesson intervention delivered on a rolling basis between March 2015 and May 2016. Year 4 (aged 8-9) classes in primary schools in the Greater London region were approached to participate in the VT study. Schools were recruited by contact with local Public Health and School Sport Partnership organisations and through enquiries elicited from the study website (www.virtualtravellerstudy.wordpress.com). One Year 4 class in each of the ten recruited schools was informed about the project by the lead author, with informed consent signed by parents/carers received from 87.1% (n=264/303; Figure 1) of pupils. Non-consenting pupils participated in Virtual Traveller (VT) or comparison (COM) sessions with their class but no data was collected from them.

Following initial recruitment, all participants completed baseline assessments (T0). Classes were then randomised to intervention (VT; 5 classes) or comparison (COM; 5 classes) groups via computer programme. Measures were repeated at the second (T1) and fourth week (T2) of the 6-week intervention period and at one week- (T3) and three months post-intervention (T4). COM classes received typical teaching, with the full VT programme supplied to use at the end of the study period (waiting list control). Ethical approval was granted by the XXXXXX Research Ethics Committee (Ref: 3500-004).

Intervention

Virtual Traveller (VT) was a programme of pre-prepared physically active lesson sessions, developed following feasibility work (Reference blinded). It consisted of 3 x 10-minute physically active VFTs a week over a 6-week period (18 sessions in total). VT was designed to be integrated into year 4 (8-9 years) National Curriculum maths and English teaching (Department for Education, 2013) and was developed with consultation from teachers with recent Year 4 teaching experience (Reference blinded). After an initial 30-minute training session, VT was provided as Powerpoint sessions via USB stick, to be delivered by teachers on existing classroom interactive whiteboards. COM teachers received this training after study data collection.

Sessions included embedded Google Earth videos showing transitions between different global locations. Accompanying text provided questions on session content and prompted children to simulate appropriate on-the-spot movements of moderate-to-vigorous intensity as they 'travelled' to- and interacted with locations. For example, children ran on-the-spot as they travelled between London and New York City when learning about explanation texts, before performing jumping jacks or high kicks to show whether quiz questions on the topic were true or false (Session E4: Explanation texts). Students stood behind their desks to complete these movements. Behaviour Change Techniques from the Behaviour Change Technique Taxonomy version 1 (BCTTv1) (Michie et al., 2013) were embedded throughout teacher training and the intervention itself (Reference blinded). For example, goal-setting (BCT 1.1) was used during teacher training where teachers agreed to deliver three VT sessions a week. An overview of the whole VT programme, detailed descriptions of example

maths and English sessions and the BCTs used can be seen in the study protocol (Reference blinded).

Measures

- Demographic measures

Pupil and teacher demographics were assessed by questionnaire at baseline (T0). Weight was assessed at baseline to the nearest 0.1 kg (Weight Watchers 8961U electronic scales, Milton Keynes, UK) and height to the nearest mm (2 metre tape measure) to calculate Body Mass Index (BMI; kg÷ m^2). Underweight, overweight and obesity prevalence was estimated using the 2nd, 85th and 95th percentiles of the 1990 UK reference curves (Cole, Freeman, & Preece, 1995).

- Outcome measures

Outcome measures to assess the effectiveness of Virtual Traveller were assessed at T0 (baseline), at weeks 2 (T1) and 4 (T2) of the six-week intervention and at one week (T3) and three months (T4) post-intervention. Primary outcome measures were sedentary behaviour (SB), light physical activity (LPA) and moderate-to-vigorous physical activity (MVPA) during school and weekend-days. Secondary outcome measures were SB, LPA and MVPA during lessons, on-task behaviour and student engagement. All data collection was administered by trained researchers, un-blinded to classes' allocation to VT and COM groups (Reference blinded).

Physical activity outcomes were assessed using Actigraph GT1M accelerometers, shown to be highly valid and reliable in children (Kim, Beets, & Welk, 2012). At each data collection

phase, accelerometers were worn for four consecutive days including two school- and two weekend-days. A device was attached to each participant's waist on their right hip with an adjustable elastic strap. Accelerometers were activated at 09:00 on Day 1 when accelerometers were distributed at the start of school and de-activated at 23:59 on Day 4. This provided a total of 86 hours maximum wear time for each data collection phase (Reference blinded). A valid accelerometer day was defined as at least 500 minutes wear time between 07:00 and 00:00 (Ekelund, Luan, Sherar, & et al., 2012). Participants were included in the analysis if they provided at last three days of valid accelerometer wear time (including one VT day in intervention pupils; Figure 1). Data was collected in 5-second epochs (Cain, Sallis, Conway, Van Dyck, & Calhoon, 2013) and analysed using Pulsford cut-points (Pulsford et al., 2011) to classify activity as sedentary: (<100 CPM), light (100-2240 CPM), moderate (2241-3840 CPM) or vigorous (≥3841 CPM). Non-wear was defined as 60 minutes of consecutive zeros (Troiano et al., 2008). Using all valid days, a daily average for time in SB, LPA and MVPA was calculated in minutes per day. Raw data was extracted from each Actigraph and analysed using ActiLife software (Actigraph, LLC, Fort Walton Beach, Florida).

Lesson physical activity outcomes were assessed firstly via 20-minute accelerometry assessments of VT and COM sessions and also via 20-minute observed assessments using the well-validated Children's Activity Rating Scale (CARS) (Finn & Specker, 2000; Puhl, Greaves, Hoyt, & Baranowski, 1990). Participating pupils were observed in turn for 4 seconds (Merrett & Wheldall, 1986) using a pre-recorded audio file during VT and COM lessons, with data recorded on a standardised score sheet. Pupils' movements were rated from 1 (stationary) to 5 (fast movement) across the observation period to provide a mean

score for each individual (Reference blinded; Puhl et al., 1990). One session in each participating class (n=10 sessions, 20% of all sessions observed) was observed by two researchers to allow reliability assessments. Inter-rater reliability across all CARS observations was high (ICC = 0.75) (Cicchetti, 1994). On-task behaviour was assessed simultaneously alongside CARS observation using the Observing Teachers and Pupils in Classrooms (OPTIC) tool (Merrett & Wheldall, 1986): well-validated within education research (Robertson & Dunsmuir, 2013). Pupil's on-task behaviour was rated as either 1 (on-task: making eye contact with teacher, following teacher's instructions etc) or 2 (off-task). Inter-rater reliability across all OPTIC observations was good (ICC = 0.66) (Cicchetti, 1994). Student engagement was assessed using the pupil-completed Student Engagement Instrument – Elementary version (SEI-E) questionnaire (Carter, Reschly, Lovelace, Appleton, & Thompson, 2012): a recent adaption of the well-validated Student Engagement Instrument (Appleton, Christenson, Kim, & Reschly, 2006) to primary school-aged children. The SEI-E features 24 items and assesses four constructs: Teacher-Student Relationships (TSR; 9 items), Peer Support for Learning (PSL; 6 items), Future Goals and Aspirations (FGA; 5 items) and Family Support for Learning (FSL; 4 items) (Carter et al., 2012). All items are 4-point Likert scales and the questionnaire takes 15-20 minutes to complete. A full process evaluation of the VT intervention was also performed (Reference blinded), to be reported in a subsequent paper.

Data analysis

Independent *t*-tests comparing VT and COM groups were performed for each outcome and assessment period. As the SEI-E (Carter et al., 2012) has not yet been tested in a UK sample, Principal Components Analysis was used to assess its structure across all completed

questionnaires (Online Supplementary Material) using SPSS for Windows (Version 19.0). This study was a cluster-randomised controlled trial, with randomisation to intervention groups done by class rather than individual pupils. Multilevel modelling was hence used to reflect the hierarchical relationships between assessment point, pupils and classes (Campbell, Mollison, Steen, Grimshaw, & Eccles, 2000). *A priori* sample size analysis was run to reflect this analysis (Maas & Hox, 2005), with calculations based on baseline post-test correlation scores of r=0.30 (Riley et al., 2016), 80% power, *α* levels set at $p<0.05$, an intraclass correlation coefficient of ICC=0.15 and a maximum number of classes of J=10, with n=140 required overall (Reference blinded). With n=219 in the analytic sample, this study hence exceeded this minimum sample size requirement.

Multilevel regression analyses were conducted using Stata (Version 12.0), with analyses performed in accordance with past physically active lesson intervention studies (de Greeff et al., 2016; Mullender-Wijnsma et al., 2015). Three-level models were constructed, with measurements at each time-point (level 1) nested within individual pupils (level 2) nested within classes (level 3). Random intercept models were developed to assess the differences between levels in impact of intervention (Virtual Traveller or control) and time-point (baseline (T0), during (T1 & T2) and post-test (T3 & T4)) and the group-by-time interaction. Outcomes at T4 were used as the dependent variables, with three models for each outcome built to investigate the effects of the intervention. The covariates model contained sex, ethnicity (white pupils coded as 0 and non-white pupils coded as 1) and measurement period (categorical: comparing scores of baseline (T0) with the intervention periods (T1 & T2) and follow-up periods (T3, T4)) as fixed effects. Model 1 added condition as a fixed effect: to investigate whether the intervention group differed from the control group.

Model 2 contained Model 1 and condition x measurement period interactions as additional fixed effects. Results of Model 2 are presented in all reporting and tables to show the most adjusted version of analysis. The model fit was evaluated by comparing the deviance of the covariates model with the deviance of Models 1 and 2. Alpha levels were set at $p<0.05$.

Results

Ten Year 4 (aged 8-9) classes from ten different primary schools were recruited to the study. Of the initial 264 pupils that were recruited (Figure 1), 133 (5 schools) were allocated to the VT intervention group and 131 (5 schools) were allocated to the COM group. No classes dropped out during the study. A total of 219 pupils (83.0% of those recruited) provided valid data in at least one measurement period and were included in the analytic sample (Table 1). At T0, 211 pupils produced valid data for at least one outcome variable, falling to 209 pupils at T4 (three month follow-up; 79.2% of recruited pupils; Figure 1). Absenteeism and no longer wanting to participate were common reasons for attrition, with participants able to re-enter the study at later data collection points. 50.7% of the analytic sample were male, with 52.1% from ethnic minority groups and 30.6% from low household income backgrounds (<£15,000; Table 1). There were no significant differences in demographic variables between VT and COM groups (Table 1).

Table 2 presents pupils' mean scores of physical activity outcomes. No intervention effects were seen for the primary study outcomes of school and weekend day SB, school day LPA and weekend day MVPA (Table 3). However for the remaining primary study outcomes, multilevel modelling analysis found higher school-day MVPA in the VT group at T1 only (first

intervention period: 60.8 minutes (SD=8.31) in VT group vs 56.1 minutes (SD=10.38) in COM group; *B*= 6.02 (1.90); 95% CI, 2.30, 9.74; *p*<0.01; Table 3), with no differences at either follow-up period. Also, a significant difference between intervention groups was observed for weekend-day LPA at T3 only (one week follow-up period: 49.6 minutes (SD=9.66) in VT group vs 47.2 minutes (10.52) in control group; *B*=10.33 (5.17); 95% CI, 0.21, 20.46; *p*=0.045), showing greater LPA in the VT group than the COM group (Table 3).

All lesson-time physical activity outcomes showed significant differences between study groups during the intervention period (T1 & T2; Table 3), with the VT group demonstrating significantly less accelerometer-assessed SB, more LPA and MVPA, as well as greater observed activity assessed with the CARS tool. Overall, VT lessons contributed 3.6% (SD=1.91) of daily MVPA compared to 0.5% (SD=0.57) in COM lessons. There were no significant differences in VT pupils' activity levels during the intervention (T1 & T2). Maintained effects of the intervention were not seen for any lesson physical activity outcome at either follow-up period (T3 & T4).

Table 2 presents pupils' mean scores of on-task behaviour and student engagement outcomes. Multilevel modelling analysis found significantly higher on-task behaviour in the VT compared to COM group at both intervention points (T1: 1.86/2 (SD=0.06) in VT group vs 1.77 (SD=0.07) in COM group; *B*=0.08 (0.01); 95% CI, 0.06, 0.11; *p*<0.001; T2: 1.85/2 (SD=0.08) in VT group vs 1.76 (SD=0.06) in COM group; *B*=0.09 (0.01); 95% CI, 0.06, 0.11; *p*<0.001)(Table 4). There were no significant differences in VT pupils' on-task behaviour during the intervention (T1 & T2). However this intervention group difference was not maintained at either follow-up period (T3 & T4). No differences in any SEI-E student engagement outcomes were observed at any time-point (Tables 2 & 4).

Discussion

The results of this study are a unique contribution to the literature on physically active lesson interventions in various ways. Firstly, by assessing physical activity across full days with accelerometry it was shown that Virtual Traveller did not have any clear effect on overall school and weekend activity levels, rejecting Hypothesis 1. Significantly greater school-day MVPA in the VT group was seen at T1 only, although the difference was small. This contrasts with previous results showing physically active teaching to have effects on school time activity (Donnelly et al., 2009; Norris et al., 2015a) and also with the only previous study to assess weekend activity, which found positive, sustained benefits (Donnelly et al., 2009). Virtual Traveller improved lesson-time physical activity as assessed by accelerometers and observations, confirming Hypothesis 2 and concurring with the majority of previous physically active lesson research (Norris et al., 2015a). Secondly, this study assessed activity twice during the intervention period to track any potential change with repeated session exposure. Importantly, no significant changes were seen in lesson-time SB, LPA or MVPA levels within the intervention group during the intervention (T1 & T2). This suggests that Virtual Traveller sessions did not have depreciating effects on lesson activity over time, opposing concerns from teachers in qualitative feasibility work that pupils may become less active during exposure to sessions (Reference blinded). However as previously shown, this increased lesson-time activity did not produce any significant differences in overall activity levels. As Virtual Traveller was performed using on-the-spot actions (Reference blinded), it may be that these movements did not elicit sufficiently intense activity to lead to subsequent increased overall activity.

Additionally, on-task behaviour (behavioural student engagement) was shown to be greater in the VT group during both intervention assessments (T1 & T2). This confirms Hypothesis 3 and concurs with previous physically active lesson research (Grieco et al., 2009). No reduction in mean on-task behaviour scores was seen during the Virtual Traveller intervention: suggesting sustained benefits during exposure to the sessions. Our study was the first to examine academic and cognitive student engagement in relation to physically active lessons (Norris et al., 2015a). No effects of Virtual Traveller were seen on any of the four SEI-E sub-scales. Hence although pupils' arguably experienced a novel teaching experience with Virtual Traveller (Reference blinded), this did not have any impact on pupils' cognitions surrounding learning and the school environment. This study has hence shown that physical activity can be integrated into academic lessons using existing classroom interactive whiteboards with positive (on-task behaviour) or no detrimental effects (student engagement) to educational outcomes. This extends beyond physically active lesson research finding no detrimental effects to activity with interventions not using classroom technologies (Donnelly et al., 2009; Mullender-Wijnsma et al., 2016; Norris et al., 2015a). Future work is needed to assess whether longer-term physically active lessons have effects on children's' physical activity and educational outcomes.

A limitation of this study, and indeed all physically active lesson interventions, is the lack of blinding (Norris et al., 2015a). Changes to the teaching environment are very obvious to pupils and are necessary for teachers to deliver the sessions. Also academic achievement was not assessed, due to the time and resources required to assess classroom grades and administer standardized testing. Strengths of this study were its design as a cluster-randomised controlled trial and its low attrition rate. It also featured a sample of ethnically

diverse and disadvantaged pupils, whereas most other physically active lesson interventions have predominantly featured white, middle-class participants (Neelon, Hesketh, & van Sluijs, 2016).

Conclusion

The Virtual Traveller physically active lesson intervention did not produce significant changes to school-day or weekend-day physical activity levels during the intervention or at one week or three month follow-ups. However, significantly less sedentary behaviour and more physical activity was produced during VT lessons compared with control lessons. The intervention was also associated with greater on-task behaviour but no differences to student engagement. These findings suggest that physically active lessons using existing classroom interactive whiteboards can be used to initiate activity within maths and English curriculums with positive effects (on-task behaviour) or at least no detriment (student engagement) to educational outcomes.

References

Appleton, J. J., Christenson, S. L., Kim, D., & Reschly, A. L. (2006). Measuring cognitive and psychological engagement: Validation of the Student Engagement Instrument. *Journal of School Psychology, 44*(5), 427-445.

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

Biddle, S. J., Pearson, N., Ross, G. M., & Braithwaite, R. (2010). Tracking of sedentary behaviours of young people: a systematic review. *Preventive Medicine, 51*(5), 345-351.

Cain, K. L., Sallis, J. F., Conway, T. L., Van Dyck, D., & Calhoon, L. (2013). Using Accelerometers in Youth Physical Activity Studies: A Review of Methods. *Journal of Physical Activity & Health, 10*(3), 437-450.

Campbell, M. K., Mollison, J., Steen, N., Grimshaw, J. M., & Eccles, M. (2000). Analysis of cluster randomized trials in primary care: a practical approach. *Family Practice, 17*(2), 192-196.

Carson, V., Hunter, S., Kuzik, N., Wiebe, S. A., Spence, J. C., Friedman, A., . . . Hinkley, T. (2015). Systematic review of physical activity and cognitive development in early childhood. *Journal of Science in Medicine & Sport*, *19*(7), 573-8.

Carter, C. P., Reschly, A. L., Lovelace, M. D., Appleton, J. J., & Thompson, D. (2012). Measuring student engagement among elementary students: Pilot of the Student Engagement Instrument—Elementary Version. *School Psychology Quarterly, 27*(2), 61-73.

Cesa, C. C., Sbruzzi, G., Ribeiro, R. A., Barbiero, S. M., de Oliveira Petkowicz, R., Eibel, B., . . . Pellanda, L. C. (2014). Physical activity and cardiovascular risk factors in children: meta-analysis of randomized clinical trials. *Preventive Medicine, 69*(0), 54-62.

Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological assessment, 6*(4), 284.

Cole, T. J., Freeman, J. V., & Preece, M. A. (1995). Body mass index reference curves for the UK, 1990. *Archives of Disease in Childhood, 73*(1), 25-29.

de Greeff, J., Hartman, E., Mullender-Wijnsma, M., Bosker, R., Doolaard, S., & Visscher, C. (2016). Long-term effects of physically active academic lessons on physical fitness and executive functions in primary school children. *Health Education Research*, *31*(2), 185-94

Department for Education. (2013). *The National Curriculum in England: Framework Document.* London: Department for Education.

Dobbins, M., Husson, H., DeCorby, K., & LaRocca, R. L. (2013). School-based physical activity programs for promoting physical activity and fitness in children and adolescents aged 6 to 18. *Cochrane Database Syst Rev, 2*, CD007651.

Donnelly, J. E., Greene, J. L., Gibson, C. A., Smith, B. K., Washburn, R. A., Sullivan, D. K., . . . Williams, S. L. (2009). Physical Activity Across the Curriculum (PAAC): A randomized controlled trial to promote physical activity and diminish overweight and obesity in elementary school children. *Preventive Medicine, 49*(4), 336-341.

Efrat, M. (2011). The relationship between low-income and minority children's physical activity and academic-related outcomes: a review of the literature. *Health Education & Behavior*, 38(5), 441-451.

Ekelund, U., Luan, J., Sherar, L. B., & et al. (2012). Moderate to vigorous physical activity and sedentary time and cardiometabolic risk factors in children and adolescents. *JAMA, 307*(7), 704-712.

Engelen, L., Bundy, A. C., Naughton, G., Simpson, J. M., Bauman, A., Ragen, J., . . . van der Ploeg, H. P. (2013). Increasing physical activity in young primary school children--it's child's play: a cluster randomised controlled trial. *Preventive Medicine, 56*(5), 319-325.

Fedewa, A. L., & Ahn, S. (2011). The effects of physical activity and physical fitness on children's achievement and cognitive outcomes: a meta-analysis. *Research Quarterly in Exercise & Sport, 82*(3), 521-535.

Finn, J. D., & Zimmer, K. S. (2012). Student engagement: What is it? Why does it matter? *Handbook of Research on Student Engagement* (pp. 97-131). Springer.

Finn, K. J., & Specker, B. (2000). Comparison of Actiwatch activity monitor and Children's Activity Rating Scale in children. *Medicine & Science in Sports & Exercise, 32*(10), 1794-1797.

Fredricks, J., McColskey, W., Meli, J., Montrosse, B., Mordica, J., & Mooney, K. (2011). *Measuring student engagement in upper elementary through high school: a description of 21 instruments*. Washington DC: Institute of Education Sciences.

Futuresource Consulting. (2010). *Interactive displays quarterly insight: State of the Market report, Quarter 1.* Futuresource Consulting.

Grieco, L. A., Jowers, E. M., & Bartholomew, J. B. (2009). Physically active academic lessons and time on task: the moderating effect of body mass index. *Medicine & Science in Sports & Exercise, 41*(10), 1921-1926.

Howie, E. K., & Pate, R. R. (2012). Physical activity and academic achievement in children: A historical perspective. *Journal of Sport and Health Science, 1*(3), 160-169.

Kim, Y., Beets, M. W., & Welk, G. J. (2012). Everything you wanted to know about selecting the "right" Actigraph accelerometer cut-points for youth, but…: A systematic review. *Journal of Science and Medicine in Sport, 15*(4), 311-321.

LeBlanc, A. G., Broyles, S. T., Chaput, J.-P., Leduc, G., Boyer, C., Borghese, M. M., & Tremblay, M. S. (2015). Correlates of objectively measured sedentary time and self-reported screen time in Canadian children. *International Journal of Behavioral Nutrition and Physical Activity, 12*(1), 1-12.

Maas, C. J., & Hox, J. J. (2005). Sufficient sample sizes for multilevel modeling. *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences, 1*(3), 86.

Martin, R., & Murtagh, E. M. (2015). Preliminary findings of Active Classrooms: An intervention to increase physical activity levels of primary school children during class time. *Teaching and Teacher Education, 52*, 113-127.

Medical Research Council. (2013). *Developing and evaluating complex interventions: new guidance*. London: Medical Research Council.

Merrett, F., & Wheldall, K. (1986). Observing pupils and teachers in classrooms (OPTIC): A behavioural observation schedule for use in schools. *Educational Psychology, 6*(1), 57-70.

Michie, S., Fixsen, D., Grimshaw, J. M., & Eccles, M. P. (2009). Specifying and reporting complex behaviour change interventions: the need for a scientific method. *Implementation Science, 4*(40), 1-6.

Michie, S., Richardson, M., Johnston, M., Abraham, C., Francis, J., Hardeman, W., . . . Wood, C. (2013). The Behavior Change Technique Taxonomy (v1) of 93 Hierarchically Clustered Techniques: Building an International Consensus for the Reporting of Behavior Change Interventions. *Annals of Behavioral Medicine, 46*(1), 81-95.

Mullender-Wijnsma, M. J., Hartman, E., de Greeff, J. W., Bosker, R. J., Doolaard, S., & Visscher, C. (2015). Improving academic performance of school-age children by physical activity in the classroom: 1-year program evaluation. *Journal of School Health, 85*(6), 365-371.

Mullender-Wijnsma, M. J., Hartman, E., de Greeff, J. W., Bosker, R. J., Doorlaard, S., & Visscher, C. (2015). Moderate-to-vigorous physically active academic lessons and academic engagement in children with and without a social disadvantage: a within subject experimental design. *BMC Public Health, 15*.

Mullender-Wijnsma, M. J., Hartman, E., de Greeff, J. W., Doolaard, S., Bosker, R. J., & Visscher, C. (2016). Physically Active Math and Language Lessons Improve Academic Achievement: A Cluster Randomized Controlled Trial. *Pediatrics, 137*(3), 1-9.

Naylor, P. J., Nettlefold, L., Race, D., Hoy, C., Ashe, M. C., Wharf Higgins, J., & McKay, H. A. (2015). Implementation of school based physical activity interventions: A systematic review. *Preventive Medicine, 72C*, 95-115.

Neelon, S. E. B., Hesketh, K. R., & van Sluijs, E. M. (2016). Will Physically Active Lessons Improve Academic Achievement for All or Widen the Achievement Gap? *Pediatrics*, *137*(3), e20154137.

Norris, E., Dunsmuir, S., Duke-Williams, O., Stamatakis, E., & Shelton, N. (2016). Protocol for the 'Virtual Traveller' cluster-randomised controlled trial: a behaviour change

intervention to increase physical activity in primary-school Maths and English lessons. *BMJ Open, 6*(6).

Norris, E., Hamer, M., & Stamatakis, E. (2016). Active Video Games in Schools and Effects on Physical Activity and Health: A Systematic Review. *Journal of Pediatrics*, 172, 40-46.

Norris, E., Shelton, N., Dunsmuir, S., Duke-Williams, O., & Stamatakis, E. (2015a). Physically active lessons as physical activity and educational interventions: A systematic review of methods and results. *Preventive Medicine, 72*(0), 116-125.

Norris, E., Shelton, N., Dunsmuir, S., Duke-Williams, O., & Stamatakis, E. (2015b). Teacher and pupil perspectives on the use of Virtual Field Trips as physically active lessons. *BMC Research Notes, 8*(1), 719.

Norris, E., Shelton, N., Dunsmuir, S., Duke-Williams, O., & Stamatakis, E. (2015c). Virtual Field Trips as physically active lessons for primary-school children: A pilot study. *BMC Public Health, 15*, 366.

Peng, W., Crouse, J. C., & Lin, J.-H. (2013). Using Active Video Games for Physical Activity Promotion: A Systematic Review of the Current State of Research. *Health Education & Behavior, 40*(2), 171-192.

Puhl, J., Greaves, K., Hoyt, M., & Baranowski, T. (1990). Children's Activity Rating Scale (CARS): Description and Calibration. *Research Quarterly for Exercise and Sport, 61*(1), 26-36.

Pulsford, R. M., Cortina-Borja, M., Rich, C., Kinnafick, F.-E., Dezateux, C., & Griffiths, L. J. (2011). Actigraph accelerometer-defined boundaries for sedentary behaviour and physical activity intensities in 7 year old children. *PloS One, 6*(8), e21822.

Riley, N., Lubans, D. R., Holmes, K., & Morgan, P. J. (2016). Findings From the EASY Minds Cluster Randomized Controlled Trial: Evaluation of a Physical Activity Integration Program for Mathematics in Primary Schools. *Journal of Physical Activity and Health*, 13(2), 198-206.

Robertson, C., & Dunsmuir, S. (2013). Teacher stress and pupil behaviour explored through a rational-emotive behaviour therapy framework. *Educational Psychology, 33*(2), 215-232.

Schulz, K. F., Altman, D. G., & Moher, D. (2010). CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *BMC Medicine, 8*(1), 18.

Stamatakis, E., Coombs, N., Tiling, K., Mattocks, C., Cooper, A., Hardy, L. L., & Lawlor, D. A. (2015). Sedentary Time in Late Childhood and Cardiometabolic Risk in Adolescence. *Pediatric,*. 135(6), e1432-e1441.

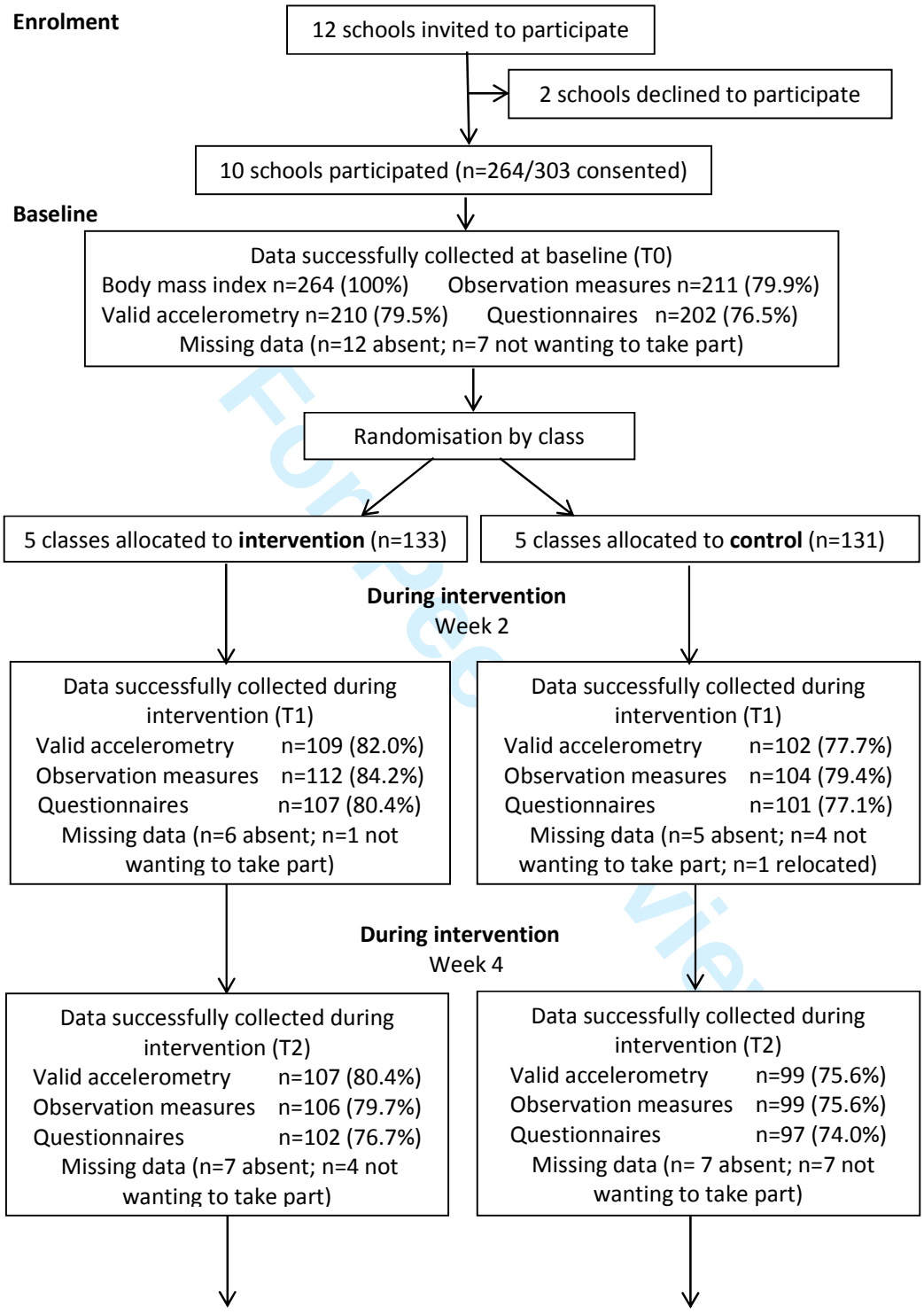
Telama, R. (2009). Tracking of physical activity from childhood to adulthood: a review. *Obesity Facts, 2*(3), 187-195.

Troiano, R. P., Berrigan, D., Dodd, K. W., Mâsse, L. C., Tilert, T., & McDowell, M. (2008). Physical activity in the United States measured by accelerometer. *Medicine and Science in Sports and Exercise, 40*(1), 181.

Turner, L., & Chaloupka, F. J. (2017). Reach and Implementation of Physical Activity Breaks and Active Lessons in Elementary School Classrooms. *Health Education & Behavior, 44*(3), 370-375.

Weiler, R., Allardyce, S., Whyte, G. P., & Stamatakis, E. (2013). Is the lack of physical activity strategy for children complicit mass child neglect? *British Journal of Sports Medicine, 48*, 1010-1013.

Figure 1. Sample flowchart



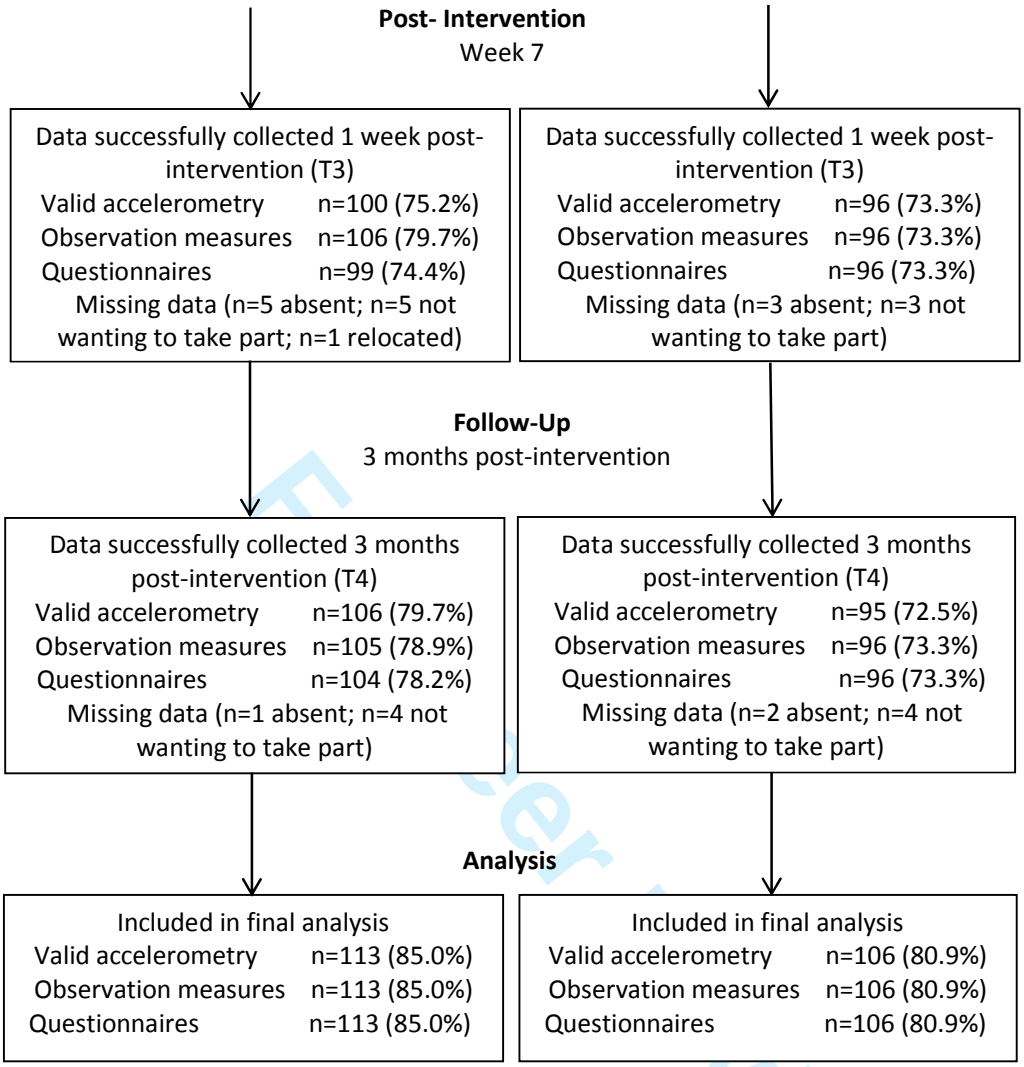


Table 1: Pupil demographics

Demographics	Overall Sample n= 219	Intervention Group n=113	Control Group n=106	p
Sex Male	n=111 (50.7%)	n=52 (46.1%)	n=59 (55.7%)	0.16
Female	n=108 (49.3%)	n=61 (54.0%)	n=47 (44.3%)	
Age Mean (SD)	8.6 (0.49)	8.6 (0.49)	8.6 (0.49)	0.88
Ethnicity				
White	n=105 (47.9%)	n=60 (53.1%)	n=45 (42.5%)	0.27
Mixed	n=15 (6.8%)	n=5 (4.4%)	n=10 (9.4%)	
Asian or Asian British	n=88 (40.2%)	n=42 (37.2%)	n=46(43.4%)	
Black or Black British	n=11 (5.0%)	n=6 (5.3%)	n=5 (4.7%)	
Chinese	n=0 (0%)	n=0 (0%)	n=0 (0%)	
Born in UK	n=167 (76.3%)	n=89 (78.8%)	n=78 (73.6%)	0.37
English as first language	n=170 (77.6%)	n=88 (77.9%)	n=82 (77.4%)	0.93
BMI Category				
Underweight	n=3 (1.4%)	n=2 (1.8%)	n=1 (0.9%)	0.99
Normal	n=134 (61.2%)	n=68 (60.2%)	n=66 (62.3%)	
Overweight	n=66 (30.1%)	n=35 (31.0%)	n=31 (29.2%)	
Obese	n=16 (7.3%)	n=8 (7.1%)	n=8 (7.5%)	
Special Educational Needs	n=3 (1.4%)	n=2 (1.8%)	n=1 (0.9%)	0.60
Physical difficulties	n=3 (1.4%)	n=1 (0.9%)	n=2 (1.9%)	0.53
Free School Meals	n=50 (22.8%)	n=28 (24.8%)	n=22 (20.8%)	0.48
Total household income				
Under £15,000	n=67 (30.6%)	n=33 (29.2%)	n=34 (32.1%)	0.47
£15,000-£19,999	n=82 (37.4%)	n=47 (41.6%)	n=35 (33.0%)	
£20,000-£29,999	n=61 (27.9%)	n=31 (27.4%)	n=30 (28.3%)	
£30,000-£39,999	n=8 (3.7%)	n=2 (1.8%)	n=6 (5.7%)	
£40,000-£49,999	n=1 (0.5%)	n=0 (0%)	n=1 (0.9%)	

Notes. Independent t-tests found no significant differences of any demographic variables between intervention groups.

Table 2: Outcome scores at all time-points

	T0		T1		T2		T3		T4	
	Mean (SD)	n	Mean (SD)	n	Mean (SD)	n	Mean (SD)	n	Mean (SD)	n
Physical activity outcomes										
Primary outcomes										
School day MVPA (mins)										
Intervention	60.6 (10.26)	108	60.8 (8.31)***	105	59.0 (10.03)	103	59.4 (9.04)	99	58.8 (7.03)	101
Control	62.0 (13.27)	96	56.1 (10.38)	99	58.3 (11.04)	98	59.9 (9.88)	94	58.6 (6.53)	92
Weekend day MVPA (mins)										
Intervention	51.1 (18.69)	89	49.6 (9.66)	77	47.7 (11.46)	83	50.1 (9.03)	75	49.5 (9.36)	84
Control	49.9 (12.03)	82	47.2 (10.52)	83	50.3 (13.33)	72	49.1 (9.70)	74	50.2 (9.09)	71
Secondary outcomes										
School day SB (mins)										
Intervention	654.8 (43.79)	108	652.6 (42.19)	105	647.8 (46.04)	103	654.4 (34.31)	99	651.5 (29.12)	101
Control	647.4 (39.32)	96	654.2 (43.20)	99	647.5 (45.59)	98	648.1 (45.15)	94	649.6 (30.58)	92
School day LPA (mins)										
Intervention	145.1 (24.77)	108	139.2 (24.98)	105	143.2 (22.31)	103	144.1 (19.77)	99	137.9 (11.98)*	101
Control	149.4 (27.43)	96	141.1 (26.42)	99	145.4 (25.84)	98	149.0 (37.40)	94	144.6 (24.18)	92
Weekend day SB (mins)										
Intervention	633.1 (58.77)	89	638.9 (41.68)	77	630.8 (38.61)	83	636.6 (52.63)	75	639.6 (53.34)	84
Control	645.3 (51.74)	82	641.4 (44.99)	83	641.6 (36.51)	72	627.9 (76.56)	74	638.6 (51.16)	71
Weekend day LPA (mins)										
Intervention	128.9 (30.08)	89	121.6 (29.12)	77	120.9 (18.61)	83	119.3 (16.62)	75	116.5 (14.28)	84
Control	134.1 (28.94)	82	122.4 (32.71)	83	129.5 (35.15)	72	115.1 (18.37)	74	117.8 (14.08)	71
Lesson SB (mins)										
Intervention	16.4 (1.28)	108	10.3 (1.86)***	107	10.0 (1.75)***	104	15.6 (2.52)*	99	16.3 (1.37)	101
Control	16.5 (1.31)	96	16.3 (1.56)	99	16.6 (1.42)	98	16.4 (1.36)	93	16.6 (1.20)	92
Lesson LPA (mins)										
Intervention	3.4 (1.17)	108	7.7 (1.50)***	107	7.7 (1.39)***	104	3.6 (1.81)	99	3.4 (1.33)	101
Control	3.2 (1.23)	96	3.5 (1.43)	99	3.1 (1.28)	98	3.4 (1.25)	93	3.2 (1.30)	92
Lesson MVPA (mins)										
Intervention	0.3 (0.31)	108	1.9 (1.14)***	107	2.3 (0.98)***	104	0.7 (0.97)***	99	0.3 (0.31)	101
Control	0.26 (0.31)	96	0.22 (0.29)	99	0.3 (0.32)	98	0.3 (0.28)	93	0.3 (0.27)	92
CARS Lesson PA										
Intervention	1.4 (0.10)	108	3.6 (0.22)***	107	3.6 (0.20)***	104	1.5 (0.15)	99	1.4 (0.12)	101
Control	1.4 (0.13)	96	1.4 (0.11)	99	1.4 (0.13)	98	1.42 (0.11)	93	1.4 (0.10)	92
Educational outcomes										
Secondary outcomes										
On-task behaviour										
Intervention	1.77 (0.07)	108	1.86 (0.06)***	107	1.85 (0.08)***	104	1.76 (0.07)	99	1.77 (0.07)	101
Control	1.77 (0.06)	96	1.77 (0.07)	99	1.76 (0.06)	98	1.77 (0.06)	93	1.76 (0.07)	92
Teacher-Student Relationships (TSR)										

Intervention	23.3 (2.82)	103	24.3 (2.39)**	102	23.8 (2.83)**	101	22.9 (3.26)	97	22.9 (2.89)	100
Control	23.4 (2.89)	92	23.6 (2.84)	96	22.9 (3.08)	97	22.8 (2.84)	93	22.9 (2.91)	92
Peer Support for Learning (PSL)										
Intervention	20.4 (3.23)	103	20.6 (3.03)	102	20.7 (2.97)	101	20.4 (3.05)	97	20.4 (2.87)	100
Control	20.3 (3.85)	92	20.3 (2.98)	96	20.3 (2.91)	97	20.3 (2.95)	93	20.3 (2.66)	92
Future Goals & Aspirations (FGA)										
Intervention	16.9 (2.88)	103	16.9 (2.80)	102	16.8 (2.77)	101	16.8 (2.81)	97	16.9 (2.75)	100
Control	16.8 (2.74)	92	16.7 (2.66)	96	16.7 (2.53)	97	16.7 (2.58)	93	16.8 (2.46)	92
Family Support for Learning (FSL)										
Intervention	13.5 (1.91)	103	13.6 (1.81)	102	13.6 (1.75)	101	13.5 (1.92)	97	13.4 (1.88)	100
Control	13.6 (1.89)	92	13.5 (1.86)	96	13.5 (1.82)	97	13.4 (1.83)	93	13.5 (1.76)	92

Notes. * $p<0.05$; ** $p<0.01$; *** $p\leq0.001$; All physical activity outcomes reported in minutes except CARS; CARS stands for Children's Activity Rating Scale, rated between as between stationary (1) and fast movement (5); Lesson time is a 20-minute period; TSR, PSL, FSL & FGA are all constructs from the Student Engagement Instrument-Elementary Version (SEI-E); TSR stands for Teacher-Student Relationship (maximum score of 28); PSL stands for Peer Support for Learning (maximum score of 24); FGA stands for Future Goals and Aspirations (maximum score of 20); FSL stands for Family Support for Learning (maximum score of 16); OPTIC stands for the Observing Pupils and Teachers in the Classroom tool assessing on-task behaviour, with behaviour rated overall during 20-minute lessons as between off-task (1) or on-task (2).

Table 3. Multilevel modelling predicting three month follow-up (T4) scores for physical activity outcomes

	School day SB (mins)	School day LPA (mins)	School day MVPA (mins)	Weekend day SB (mins)	Weekend day LPA (mins)	Weekend day MVPA (mins)	Lesson SB (mins)	Lesson LPA (mins)	Lesson MVPA (mins)	CARS
Fixed effects (SE)										
Intercept	652.59 (4.84)***	152.63 (8.20)***	60.66 (1.23)***	642.43 (9.03)***	134.32 (5.22)***	47.66 (1.44)***	16.47 (0.21)***	3.18 (0.16)***	0.34 (0.08)***	1.44 (0.03)***
Sex^{x}	-3.17 (2.59)	-1.38 (1.25)	2.27 (0.62)***	2.56 (3.68)	1.47 (1.71)	3.16 (0.84)***	0.06 (0.10)	-0.04 (0.09)	-0.02 (0.04)	0.00 (0.01)
Ethnicity^	-6.27 (2.65)*	0.88 (1.33)	0.15 (0.65)	0.19 (3.90)	0.32 (1.82)	0.43 (0.84)	0.03 (0.11)	0.08 (0.09)	-0.10 (0.04)*	-0.02 (0.01)*
T1^{0}	6.82 (5.75)	-9.30 (2.74)***	-5.83 (1.37)***	-4.12 (7.73)	-10.09 (3.58)**	-2.45 (1.80)	-0.25 (0.23)	0.29 (0.20)	-0.05 (0.09)	-0.01 (0.02)
T2^{0}	0.21 (5.78)	-5.63 (2.75)*	-3.57 (1.38)**	-4.08 (8.02)	-4.02 (3.72)	0.61 (1.87)	0.11 (0.23)	-0.14 (0.20)	0.01 (0.09)	-0.01 (0.02)
T3^{0}	0.74 (5.82)	-1.23 (2.78)	-1.97 (1.39)	-15.90 (7.96)*	-20.38 (3.69)***	-0.57 (1.86)	-0.17 (0.23)	0.18 (0.20)	-0.01 (0.09)	-0.01 (0.02)
T4^{0}	2.24 (5.86)	-5.82 (2.80)*	-3.28 (1.40)*	-5.13 (8.05)	-17.70 (3.73)***	0.52 (1.88)	0.07 (0.23)	-0.02 (0.20)	-0.01 (0.09)	-0.01 (0.02)
Intervention	6.14 (6.06)	-5.24 (11.50)	-0.98 (1.57)	-12.12 (11.99)	-5.50 (7.09)	1.92 (1.78)	-0.14 (0.26)	0.17 (0.21)	-0.04 (0.10)	-0.03 (0.04)
T1*Intervention	-9.19 (7.94)	3.58 (3.80)	6.02 (1.90)**	7.49 (10.94)	3.81 (5.06)	0.70 (2.55)	-5.86 (0.31)***	4.06 (0.27)***	1.74 (0.12)***	2.24 (0.03)***
T2*Intervention	-6.97 (7.99)	3.43 (3.81)	1.96 (1.91)	0.85 (11.03)	-3.73 (5.11)	-4.07 (2.57)	-6.45 (0.32)***	4.43 (0.27)***	2.02 (0.12)***	2.20 (0.03)***
T3*Intervention	-1.05 (8.06)	0.45 (3.85)	0.68 (1.93)	19.70 (11.14)	10.33 (5.17)*	-0.76 (2.60)	-0.55 (0.32)	0.08 (0.28)	0.44 (0.12)***	0.05 (0.03)
T4*Intervention	-5.40 (8.06)	-1.30 (3.85)	1.45 (1.93)	11.48 (11.04)	5.38 (5.12)	-2.31 (2.58)	-0.11 (0.32)	0.02 (0.27)	0.05 (0.12)	0.02 (0.03)
Random effects (SE)										
Variance between classes	3.46 (2.25)	17.67 (4.02)	1.27 (0.52)	14.60 (3.95)	9.71 (2.38)	1.16 (8.07)	0.21 (0.07)	0.12 (0.06)	0.07 (0.29)	0.05 (0.01)
Variance within classes	40.02 (0.90)	19.15 (0.43)	9.57 (0.22)	49.60 (1.26)	22.99 (0.58)	11.58 (0.29)	1.59 (0.04)	1.37 (0.03)	0.62 (0.01)	0.14 (0.01)
Model deviance	-5075.38	-4371.55	-3664.54	-4215.32	-3615.78	-3055.77	-1883.09	-1728.75	-938.65	562.53

Notes. SB = sedentary time, LPA = light physical activity, MVPA = moderate-to-vigorous physical activity, CARS = Children's Activity Rating Scale; ^{x} where boys coded as 0 and girls coded as 1, ^ where white pupils coded as 0 and non-white pupils coded as 1; ^{0} indicates comparison of scores between given time-point and T0 (baseline); *B* co-efficients presented, with Standard Error (SE) in brackets; * $p<0.05$; ** $p\leq0.01$; *** $p\leq0.001$.

Table 4. Multilevel modelling predicting three month follow-up (T4) scores for on-task behaviour and student engagement outcomes

	On-task behaviour	Teacher-Student Relationships (TSR) [a]	Peer Support for Learning (PSL) [a]	Future Goals & Aspirations (FGA) [a]	Family Support for Learning (FSL) [a]
Fixed effects (SE)					
Intercept	1.77 (0.01)***	23.76 (0.34)***	19.49 (0.34)***	17.22 (0.31)***	13.92 (0.23)***
Sex[x]	0.00 (0.00)	-0.70 (0.18)***	0.84 (0.19)***	-0.89 (0.17)***	-0.15 (0.11)
Ethnicity^	0.00 (0.00)	0.06 (0.19)	0.63 (0.19)***	0.17 (0.17)	-0.49 (0.12)***
T1[0]	0.00 (0.01)	0.14 (0.41)	0.03 (0.43)	-0.16 (0.39)	-0.03 (0.26)
T2[0]	-0.01 (0.01)	-0.52 (0.41)	-0.04 (0.42)	-0.18 (0.39)	-0.08 (0.26)
T3[0]	0.00 (0.01)	-0.58 (0.42)	0.02 (0.43)	-0.17 (0.39)	-0.19 (0.27)
T4[0]	0.00 (0.01)	-0.55 (0.42)	-0.01 (0.43)	-0.07 (0.39)	-0.08 (0.27)
Intervention	0.00 (0.01)	-0.18 (0.42)	0.30 (0.42)	-0.02 (0.38)	-0.19 (0.29)
T1*Intervention	0.08 (0.01)***	0.88 (0.57)	0.14 (0.59)	0.13 (0.54)	0.15 (0.36)
T2*Intervention	0.09 (0.01)***	1.00 (0.57)	0.26 (0.59)	0.13 (0.54)	0.24 (0.36)
T3*Intervention	-0.01 (0.01)	0.15 (0.60)	-0.15 (0.59)	0.15 (0.54)	0.27 (0.37)
T4*Intervention	0.00 (0.01)	0.20 (0.58)	-0.12 (0.59)	0.09 (0.54)	0.08 (0.37)
Random effects (SE)					
Variance between classes	0.01 (0.01)	0.18 (0.14)	0.06 (0.31)	4.78 (3.58)	0.20 (0.08)
Variance within classes	0.07 (0.01)	2.83 (0.06)	2.91 (0.07)	2.65 (0.06)	1.81 (0.04)
Model deviance	1289.22	-2397.22	-2421.12	-2330.04	-1959.45

Notes. [a] indicates sub-scale from the SEI-E = Student Engagement Instrument – Elementary Version; [x] where boys coded as 0 and girls coded as 1, ^ where white pupils coded as 0 and non-white pupils coded as 1; [0] indicates comparison of scores between given time-point and T0 (baseline); *B* co-efficients presented, with Standard Error (SE) in brackets; * $p<0.05$; ** $p\leq0.01$; *** $p\leq0.001$.

Online Supplementary Material. Principal Components Analysis of Student Engagement Instrument – Elementary Version (SEI-E)

The SEI-E(Carter, Reschly, Lovelace, Appleton, & Thompson, 2012) has not yet been tested in a UK primary-school sample. Principal Components Analysis (PCA) was hence used to identify the composite sub-scale scores in this sample. Direct Oblimin rotation was used to allow inter-correlation among factors, with pattern matrix values presented to show the unique contribution of items to SEI-E factors.(Graham, Guthrie, & Thompson, 2003) Factors with eigenvalues over 1.0 were included, (Kaiser, 1960) with only item factor loadings over 0.4 considered.(Richman, 1986)

A four-factor solution was identified (factors with eigenvalue ≥1), explaining 60.54% of the cumulative variance (Additional Table 1). Only two items did not meet the minimum pattern matrix item factor loading criteria (>0.4) in their original accompanying SEI-E sub-scale. These were Teacher-Student Relationships (TSR) item 4: 'My teachers are there for me when I need them' and TSR item 6: 'My teachers are honest with me' (Table 8-4). These two items were both removed to produce a seven-item TSR scale in subsequent analysis. The factor labels proposed in the original SEI-E paper(Carter et al., 2012) were hence retained in this study, with twenty-two out of twenty-four original items retained in subsequent analysis.

Internal reliability of post-PCA SEI-E sub-scales was assessed with Cronbach's *α* (0.7-0.8: "acceptable", 0.8-0.9: "good", 0.9-1.0: 'excellent').(Kilne, 1999) Overall across all time-points, Peer Support for Learning (PSL) and Future Goals and Aspirations (FGA) sub-scales

were found to have excellent internal reliability (α=≥0.70), whilst Teacher-Student Relationships (TSR) and Family Support for Learning (FSL) sub- scales had good internal reliability (α=0.60-0.69; Additional Table 2).

Additional Table 1. Factors loadings for Student Engagement Instrument – Elementary Version (SEI-E) items from Pattern matrices (n=219)

	Factors determined through Principal Components Analysis			
Item *	**TSR**	**PSL**	**FGA**	**FSL**
Adults at my school are fair towards students most of the time (TSR1)	**0.71**			
Adults at my school listen to the students (TSR2)	**0.68**			
Teachers at my school care about students (TSR3)	**0.71**			
My teachers are there for me when I need them (TSR4)	0.36			
The rules at my school are fair (TSR5)	**0.66**			
My teachers are honest with me (TSR6)	0.32			
I like talking to the teachers here (TSR7)	**0.79**			
I feel safe at school (TSR8)	**0.53**			
Most teachers care about me as a person, not just as a student (TSR9)	**0.61**			
Other students care about me (PSL1)		**0.73**		
Students at my school are there for me when I need them (PSL2)		**0.80**		
Other students here like me the way I am (PSL3)		**0.70**		
I enjoy talking to the students here (PSL4)		**0.84**		
Students here respect what I have to say (PSL5)		**0.72**		
I have friends at school (PSL6)		**0.64**		
I plan to go to university after I finish secondary school (FGA1)			**0.82**	
Continuing to learn after secondary school is important (FGA2)			**0.66**	
School is important for reaching my future career goals (FGA3)			**0.81**	
My education will create many chances for me to reach my future goals (FGA4)			**0.82**	I
am hopeful about my future (FGA5)			**0.49**	
My family/guardian(s) are there for me when I need them (FSL1)				**0.64**
When I have problems at my school my family/guardian(s) are ready to help me (FSL2)				**0.74** My
family/guardian(s) want to know when something good happens at school (FSL3)				**0.70** My
family/guardian(s) want me to keep trying when things are tough at school (FSL4)				**0.68**

Notes: TSR, PSL, FSL & FGA are all constructs from the Student Engagement Instrument-Elementary Version (SEI-E); TSR stands for Teacher-Student Relationships, PSL stands for Peer Support for Learning, FGA stands for Future Goals and Aspirations, FSL stands for Family Support for Learning; Items denote pattern matrix loadings, with factor loadings over 0.4 in bold. *Brackets denote the original questionnaire coding of the SEI-E item.

Additional Table 2. Internal reliability of post-PCA SEI-E sub-scales

Time-point	n	Construct	Number of items	Cronbach's *α*	Skewness	Kurtosis	Shapiro-Wilk
T0	195	Teacher-Student Relationships (TSR)	7	0.62	-0.70	0.42	0.95***
		Peer Support for Learning (PSL)	6	0.86	-0.83	0.32	0.89***
		Future Goals and Aspirations (FGA)	5	0.80	-0.83	0.31	0.91***
		Family Support for Learning (FSL)	4	0.67	-0.81	0.69	0.92***
T1	198	Teacher-Student Relationships (TSR)	7	0.65	-0.79	0.79	0.94***
		Peer Support for Learning (PSL)	6	0.84	-0.78	0.17	0.89***
		Future Goals and Aspirations (FGA)	5	0.80	-0.83	0.31	0.91***
		Family Support for Learning (FSL)	4	0.65	-0.84	0.78	0.92***
T2	198	Teacher-Student Relationships (TSR)	7	0.63	-0.66	0.18	0.95***
		Peer Support for Learning (PSL)	6	0.76	-0.78	0.17	0.89***
		Future Goals and Aspirations (FGA)	5	0.75	-0.83	0.31	0.91***
		Family Support for Learning (FSL)	4	0.71	-0.84	0.78	0.92***
T3	190	Teacher-Student Relationships (TSR)	7	0.63	-0.61	-0.22	0.95***
		Peer Support for Learning (PSL)	6	0.79	-0.81	0.36	0.89***
		Future Goals and Aspirations (FGA)	5	0.75	-0.85	0.38	0.91***
		Family Support for Learning (FSL)	4	0.62	-0.84	0.88	0.92***
T4	192	Teacher-Student Relationships (TSR)	7	0.56	-0.74	0.59	0.95***
		Peer Support for Learning (PSL)	6	0.78	-0.76	0.31	0.91***
		Future Goals and Aspirations (FGA)	5	0.80	-0.83	0.37	0.91***
		Family Support for Learning (FSL)	4	0.61	-0.79	0.78	0.92***
Overall	219	Teacher-Student Relationships (TSR)	35	0.61	-0.74	0.43	0.86***
		Peer Support for Learning (PSL)	30	0.79	-0.80	0.23	0.89***
		Future Goals and Aspirations (FGA)	25	0.76	-0.83	0.35	0.90***
		Family Support for Learning (FSL)	20	0.67	-0.82	0.82	0.92***

Note: *** $p < 0.001$

Online appendix references

Carter, C. P., Reschly, A. L., Lovelace, M. D., Appleton, J. J., & Thompson, D. (2012). Measuring student engagement among elementary students: Pilot of the Student Engagement Instrument—Elementary Version. *School Psychology Quarterly, 27*(2), 61-73.

Graham, J. M., Guthrie, A. C., & Thompson, B. (2003). Consequences of not interpreting structure coefficients in published CFA research: A reminder. *Structural Equation Modeling, 10*(1), 142-153.

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement*, *20*(1), 141-151.

Kilne, P. (1999). *The Handbook of Psychological Testing* (Vol. 2nd Edition.). London: Routledge.

Richman, M. B. (1986). Rotation of principal components. *Journal of climatology, 6*(3), 293-335.