Elsevier Editorial System(tm) for Journal of Science and Medicine in Sport

Manuscript Draft

Manuscript Number: JSAMS-D-17-00407R1

Title: Validation of an observation tool to assess physical activity-promoting physical education lessons in high schools: SOFIT+

Article Type: Original Research

Keywords: Adolescent, Health, Students, MVPA, Validity, Teaching

Corresponding Author: Professor Stuart J Fairclough, PhD

Corresponding Author's Institution: Edge Hill University

First Author: Stuart J Fairclough, PhD

Order of Authors: Stuart J Fairclough, PhD; R. Glenn Weaver, PhD; Siobhan Johnson, BA; Jack Rawlinson, BA

Abstract: Objectives. SOFIT+ is an observation tool to measure teacher practices related to moderate-to-vigorous physical activity (MVPA) promotion during physical education (PE). The objective of the study was to examine the validity of SOFIT+ during high school PE lessons.
Design. This cross-sectional, observational study tested the construct validity of SOFIT+ in boys' and girls' high school PE lessons.
Methods. Twenty-one PE lessons were video-recorded and retrospectively coded using SOFIT+. Students wore hip-mounted accelerometers during lessons as an objective measure of MVPA. Multinomial logistic regression was used to estimate the likelihood of students engaging in MVPA during different teacher practices represented by observed individual codes and a combined SOFIT+ index-score.
Results. Fourteen individual SOFIT+ variables demonstrated a statistically significant relationship with girls' and boys' MVPA. Observed lesson segments identified as high MVPA-promoting were related to an increased likelihood of girls engaging in 5-10 (OR = 2.86 [95%CI 2.41-3.40]), 15-25 (OR = 7.41 [95%CI 6.05-9.06]), and 30-40 (OR = 22.70 [95%CI 16.97-30.37]) seconds of MVPA. For boys, observed high-MVPA promoting segments were related to an increased likelihood of engaging in 5-10 (OR = 1.71 [95%CI 1.45-2.01]), 15-25 (OR = 2.69 [95%CI 2.31-3.13]) and 30-40 (OR = 4.26 [95%CI 3.44-5.29]) seconds of MVPA.
Conclusions. Teacher practices during high school PE lessons are significantly related to students' participation in MVPA. SOFIT+ is a valid and reliable tool to examine relationships between PE teacher practices and student MVPA during PE.

Cover Letter

Edge Hill University

Prof Stuart Fairclough,
Physical Activity and Health Research Group
Sport and Physical Activity Department
Edge Hill University
St Helens Road
Ormskirk
Lancs
L39 4QP
UK

Tel: 00 44 1695 584143
Email: stuart.fairclough@edgehill.ac.uk

25th August 2017

Dear Prof Waddington,

On behalf of my co-authors I am writing to inform you of a revised manuscript submission to the Journal of Science and Medicine in Sport, entitled *Validation of an observation tool to assess physical activity-promoting physical education lessons in high schools: SOFIT+.*

Category of article: original research.
Sub-discipline: physical activity and health
Financial support for the project: No external financial support

On behalf of the co-authors, we look forward to receiving reviewers' comments in due course.

Yours faithfully,

Prof. Stuart Fairclough

Professor of Physical Activity Education

Edge Hill University, UK

Responses to review comments

Reviewer #1:

Abstract

Perhaps consider more indicative keywords - 'odds ratio' 'promotion' seem a little vague.

- We thank the reviewer for this observation. We have amended the key words by adding ‘validity’ and ‘teaching’ in place of the keywords highlighted.

This is a clunky opening sentence that doesn't read that well. I would also remove the unnecessary adverb in this sentence.

SOFIT+ is an observation tool to comprehensively measure teacher

SOFIT+ practices that both promote and limit students' moderate-

SOFIT+ to-vigorous physical activity (MVPA) during physical education

Ditto the final sentence.

SOFIT+ produces valid and reliable data and can examine relationships between a range of physical activity-promoting and -limiting practices and student MVPA.

- The opening and final sentence of the abstract have been modified in line with the reviewer’s comments (lines 2-3 and 20-21).

Introduction:

L.28. Remove the connective adverb.

- ‘Moreover’ has been removed (line 25).

L.29 Advocated for what?

- We have clarified that school-based PA are advocated to promote health and wellbeing (line 28).

L.29-31 During adolescence I take it?

- This is a general statement that applies to all children and adolescents, so the applicability to adolescence is implicit (lines 29-32).

L.31-34 A minor point, but the word youth and adolescence occur inter-changeably. Can one term be adopted for consistency? If not, include the age range you are talking about. This happens in the sentence above.

- This is a very useful observation. To improve consistency, reference to ‘youth’ has been removed throughout the manuscript, and replaced with ‘adolescent’ or ‘adolescence’ where appropriate.

L.38 Abbreviations for US and UK. In full first time around? Plus as the word secondary school is been adopted, not high school, would it be tidier to flip the US and UK around. This happens a lot throughout the manuscript. It is a UK based study so keep the terms UK based. A decision may need to be taken on the term 'students' or 'pupils'. I have no preference either way, if whatever is included can be supported and is consistent.

- As the study was conducted in the UK we have taken the reviewer’s suggestion and flipped the reference to United Kingdom and United States (line 37). We have though retained the use of ‘students’ throughout as this term is recognised internationally and is often used in the UK context.

L.40-41 This is not clear, and needs to be re-worded.

- We have clarified the meaning here by inserting ‘i.e., minutes’ in parentheses (line 38).

L.43-44 Ref needed here to hammer home the point.

- Reference has been provided (line 40).

L. 44-45 I think I understand the point being made here, but again it is vague, could this be more precise?

- We have provided more detail in this sentence to improve the clarity. It now reads: ‘Thus, evaluation of PE quality relative to physical activity promotion requires additional indicators which complement MVPA’ (lines 40-41).

L. 45 Quite a leap to SOFIT. As this is the premise for the study consider a new para here. Also I think SOFIT does more than provide 'teacher insights'. I think a statement about its validity/reliability in previous PA research needs to be included.

- We have started the explanation of SOFIT+ as a new paragraph as suggested, and have also expanded the description of the original SOFIT instrument (lines 43-45). The 3000 word limit and 30 citation limit does not allow for a detailed overview of previous SOFIT research but we hope that readers can use the reference to the original SOFIT study as a springboard from which to locate subsequent articles using the instrument.

L.47 How was it modified? This is very vague and requires more information to convince me the modifications are valid/reliable etc. Has SOFIT+ been used in previous PE studies?

- As well as providing the original SOFIT+ citation here (#14), we have included reference to the SOFIT+ sections and referred readers to the supplementary file 1 which provides a description of each SOFIT+ code (lines 47-48). Ideally we would have provided all of this detail in the manuscript but the word limit constraints were too great to do this. Reference to the previous SOFIT+ PE study is initially made in line 47.

L.49 Best practice data is again very vague. Could we have some actual figures please?

- The SOFIT+ 'best practice' codes and definitions are described in the supplementary file 1 to which readers are referred to (line 48) and the original SOFIT+ manuscript which is referenced throughout (lines 47 and 50 in the Introduction).

L.51 Just a point here on the slippage in language between US terms and UK. For example, elementary school (US) and secondary (UK). Can secondary be replaced with High School?

- 'Elementary school' is used because the study referred to was based in that setting in the US. On reflection, we think that some readers may not be too familiar with the term 'secondary school' and so this has been changed throughout to 'high school' in line with the reviewer's suggestion.

Methods:

74-76 Is this considered the 'gold standard' for accelerometry data capture? What steps were considered to reduce error? Also as the device was attached to the hip, how were any hip-motion activities assessed? The same I guess applies to vertical measurements? Reliability status of ActiGraph? Just thinking if you are to go and compare results between different studies. What was the frequency calibration?

- Typically physical activity studies in children have used waist-mounted accelerometers as this position best approximates the centre of mass and provides accurate estimates of PA during ambulatory activities [1]. It is well established that waist-mounted accelerometers will not capture activities performed when the hip is stationary (which do occur during PE), such as upper body movements and we have alluded to this in the limitations. However, we are confident that all movement was captured in this study as none of the content observed involved non-ambulatory

activities. That is, the content observed included various team invasion games, track and field, fitness circuits, and gymnastics/dance. See lines 142-144.

In recent years the wrist has become more prominent as an accelerometer wear site [2] and this may overcome some of the challenges of capturing movements that are not dependent on ambulation. Wrist-worn monitors also have limitations though such as choice of dominant or nondominant wrist, and lack of comparability with previous studies using waist placed accelerometers [3]. Presently there is no consensus as to which wear site (if any) is the 'gold standard' [4]. We chose the waist as the wear location to be consistent with the previous SOFIT+ study protocol [5] and the majority of previous school-based studies that have used ActiGraph monitors. To reduce error, the devices were initialised and downloaded on the same PC by the same person, and the instructions and demonstrations for wear were provided by the same two researchers (one male, one female) following a standardised script. The validity and reliability of the ActiGraph GT3x+ is well established [6-8] and we used the MVPA cutpoints developed by Evenson et al. [9] which have demonstrated strong classification accuracy compared to other cutpoints [10]. The monitors were set to record at a frequency of 30 Hz (line 75).

L.77 Positioning of the camera - did this allow for the capture of the whole lesson? Was it elevated at all?

- The camera was positioned in the corner of the work space where possible. The GoPro widescreen/fisheye lense ensured that all of the teaching area was captured on the video. We have referred to this in lines 78-79.

L80 Was a specific SOFIT+ recording form used?

- The specific SOFIT+ recording form was developed and deployed using Pendragon Forms Universal© on Samsung Galaxy Tablets© (lines 80-81).

L91 I take it these were SOFIT+ training videos?

How long did it take for both trained observers to become proficient in the use of SOFIT+. I think this needs to be included.

What happened in the training when there was observer disagreement?

- The training videos were developed specifically for use with SOFIT+ and we have made this clear in the manuscript (lines 89 and 91). As indicated in the manuscript

training took approximately 14 hours for the two observers to become proficient and we have detailed this in line 93. In the event of disagreement between observers they would review the SOFIT+ codes and definitions and discuss their interpretation in the context of the section of video under consideration. Any instances such as this were flagged with the second author who developed the SOFIT+ protocol and a threeway discussion ensued which again involved reviewing the video and SOFIT+ codes. This process resulted in consensus as to which were the appropriate SOFIT+ codes.

Results:

L.152 Caps for 'Content'

- Changed as suggested (line 148).

Discussion:

L.190 hypothesised is spelt using US spelling earlier in the document. Be consistent.

- Thank you for highlighting this. The UK spelling is used throughout.

L. 208-09 This does not make sense.

- The sentence has been amended to improve the clarity (lines 205-206).

L.210-11 How?

- An explanation of how sex-specific differences may be related to delivery of the girls' and boys' lessons follows throughout the rest of the paragraph (lines 206-224). To make this link more obvious, the word 'because' has been included at the end end of the initial sentence (line 208).

L.239 Remove however.

- This has been done.

This is a well-written paper and to my knowledge the first to examine the use of SOFIT+ in a secondary UK based setting. The obvious limitation with the study is the sample size, and I feel as though this will be the deal-breaker. Two sessions were lost due to the battery malfunction which left 19 for final analysis, which for this type of study is modest at best. To be fair, the authors do allude to this as a limitation at the end of the document. The

introduction requires some tweaking and the method requires some additional information which is included in the line-by-line comments. The results/tables were well presented and included the CI levels.

- We thank the reviewer for taking the time time to read and comment on the paper. The sample size is consistent with the previous SOFIT+ article, though we acknowledge that the paper may have been strenghthened with the inclusion of more lessons. We were working within a tight data collection window that related to acces to the schools and availability of research assistants to undertake the data collection. Thus, it is unlikely that we would have been able observe significantly more lessons than we did. Further, there are two points that we would like to make regarding sample size. The activity codes used in the original SOFIT instrument were validated against the heart rate of only 19 children.[11] This is far fewer than the 225 students included in this study. A similar observation instrument created to observe child interactions during recess included 24 days of observation.[12] The sample of lessons observed in this study is of a similar size. Also, because validity the SOFIT+ variables was tested at the scan level and not the class level our sample is actually much larger than the 19 classes observed (i.e., 1079 scans which are nested within the 19 classes).

Reviewer #3: General comments.

Thank you for the opportunity to review this paper. Without reading the authors previous study (reference 14) it was difficult to provide a concise peer-review. Many of my comments/ questions could probably be addressed by reading this paper, and I understand that with space restrictions the authors need to rely on citing their previous work, however if this paper is to be stand alone, some detail that is integral to a readers understanding of the paper needs to be added. I hope my comments are useful for the authors in improving the manuscript and can be done in word limits.

- We thank the reviewer for taking the time time to read and comment on the paper.

Introduction

The paper is generally well-written. More detail around SOFIT+ is really needed to help set up the results that are presented. Pg3 Line 45-48 states "For two decades the System for Observing Fitness Instruction Time (SOFIT) has provided insight into teacher practices that

influence MVPA engagement during PE. Recently, SOFIT was modified to develop an observation tool (SOFIT+) that comprehensively measures teacher practices that both promote and limit MVPA during PE". What are these things? When I got to the results I am presented with findings for lesson and activity content that I thought would have already been validated in previous studies. Perhaps it could simply be addressed by including (if word limit restricts even as a table in supplementary material) how SOFIT+ differs from SOFIT, so readers can easily see what the new components of teacher practices are included and being assessed. It may also help to edit the aim to add a bit more detail about these practices for example "The study purpose therefore, was to examine the validity of SOFIT+ encompassing XYZ during secondary school PE lessons."

- As well as providing the original SOFIT+ citation here (#14), we have included reference to the SOFIT+ sections and referred readers to the supplementary file 1 which provides a description of each SOFIT+ code (lines 47-48). Ideally we would have provided all of this detail in the manuscript but the word limit constraints were too great to do this. Reference to the previous SOFIT+ PE study is initially made in line 47. We thank the reviewer for their suggested edit to the aim, but on reflection we prefer to keep the language as it stands to maintain clarity and succinctness.

Methods

Understand that structure of JSAMS methods does not allow for subheadings however some detail was missing for example; study design; sampling frame (how/why were these two schools selected?); setting detail (are they low/ high SES; rural/ urban); recruitment- how were schools, teachers, students recruited? was there any method for selecting which lessons were observed? Was there any eligibility criteria for lessons that were observed e.g. was a cross country lesson eligible for observation or did it have to be a teacher directed lesson? could it be a casual/ relief teacher/ external provider/coach teaching the lesson or did it have to be usual practice?

- We have inserted additional details about the study design (cross-sectional), sampling (convenience), school locations (urban and relatively medium-high SES), and teacher status (usual rather than relief; lines 65-71). All the lessons were teacher-directed (i.e., taught) and lesson selection was based on timetabled schedules of students in Years 7-9 that did not clash between the two schools.

What was the rationale for only selecting 21 lessons from 8 teachers? Power calc? What effect did only having 19 impact on this?

- No power calculation was undertaken. We aimed to collect data from at least the same number of lessons (20) as Weaver and colleagues in their original SOFIT+ study [5] and would have achieved this were it not for the technical malfunctions. The sample size is consistent with the previous SOFIT+ article, though we acknowledge that the paper may have been strenghthened with the inclusion of more lessons. We were working within a tight data collection window that related to acces to the schools and availability of research assistants to undertake the data collection. Thus, it is unlikely that we would have been able observe significantly more lessons than we did. Further, there are two points that we would like to make regarding sample size. The activity codes used in the original SOFIT instrument were validated against the heart rate of only 19 children.[11] This is far fewer than the 225 students included in this study. A similar observation instrument created to observe child interactions during recess included 24 days of observation.[12] The sample of lessons observed in this study is of a similar size. Also, because validity the SOFIT+ variables was tested at the scan level and not the class level our sample is actually much larger than the 19 classes observed (i.e., 1079 scans which are nested within the 19 classes). Eight teachers were selected because they were timetabled to deliver the designated Year 7, 8 and 9 PE lessons.

Were your methods consistent with previous studies that validated the original SOFIT? It would be good to set your study in context of those?

- We thank the reviewer for this thoughtful comment. The answer is yes and no. First, SOFIT activity codes have been validated against a variety of wearable criterion measures of activity, including accelerometers, in the past.[11, 13-22] Thus, this process is consistent with the previous literature. However, previous studies using the SOFIT instrument and other direct observation instruments have only provided the inter-rater reliability of behaviors and/or contextual variables. We have provided reliability data for these variables but have also tested the construct validity of the behaviors as related to to objectively measured physical activity. While this process has never been completed with the SOFIT instrument it has been completed for other direct observation instruments.[23] We consider this to be a step forward and a strength of this study.

Pg 6 line 97 could the authors clarify if the content validity has been established with elementary or secondary teachers? If only elementary is this true for secondary schools?

- The content validity was established using a modified Delphi method. This involved experts (academics and PE teachers with experience of teaching elementary and high school students) reviewing and evaluating a list of teacher practices that impact MVPA in PE [5]. Thus, we believe that the SOFIT+ codes are valid for both elementary and high school PE settings.

Whilst I appreciate word limits of this journal, pg 6 lines 107-111 more detail is required here rather than simply referring readers off to the previous study. This may also help with my initial comment around this SOFIT+ tool.

- We have made it clearer that the SOFIT+ index score is described fully in the previous paper, but is also outlined in brief in the current manuscript (lines 107-115).

Were any data collected on teacher characteristics i.e. years of teaching experience, gender? How did you control for these variables in analysis as they have been associated with MVPA in PE?

- We collected data on the teacher gender. The lessons were single-gender and taught by teachers of the same gender as the students. For this reason there was no adjustment in the analyses for teacher gender. No information though was collected on teacher experience. If such data were collected we would hypothesise that teachers with more experience would engage in more of the PA promoting competencies, which in turn would increase the likelihood of students engaging in PA. This is because experienced teachers are generally more likely to be practitioners who manage and teach more effectively. In this sense, experience could be viewed as a proxy for effective practices which were captured using SOFIT+. We are grateful for the reviewer flagging this point up, and it is something we would consider collecting data in future SOFIT+ work.

Results

What % of the class is represented?

- All of the students consented to take part so 100% of the students in the classes were represented.

The rationale for presenting the data in 5-10secs; 15-25 secs; 30-40 secs isn't very clear. If not a strong reason to do so, it may be more compelling to simply present the longer segments as this has greater public health relevance and applicability to use this tool at a population level.

- Presenting the data in this manner was chosen to represent the intermittent and sporadic nature of children's physical activity [24] and to improve the ability to capture the impact of the PA promoting practices on children's transitory physical activity patterns.[25, 26] Thus, we would argue that presenting the data in longer segments would be in direct conflict with the way that children move. That is, children move in short bursts not long sustained periods of activity. Further, presenting the data in this manner allowed any linear patterns to be highlighted (e.g., increased odds of MVPA engagement with increased duration of a certain context or activity). An example of this is boys' MVPA and game play, as described in Table 3.

For both boys and girls the results includes a sentence "Segments identified as high promoting by the SOFIT+ index" what are these? Apologies if I have missed this in the paper, I can see it at the bottom of each table however is this the cumulative of all practices under lesson/activity context and teacher behaviour?

- The process for calculating the SOFIT+ index score and which segments were identified as high promoting is described in the Methods in lines 108-115.

Discussion

Lines 226-228 "It is therefore likely that lesson contexts, activity contexts, and teacher behaviours combined to influence students' MVPA during Small-Sided and Large-Sided Activities". This was a question that kept coming to mind throughout the paper. How much is it a combination of all/ some practices together? Can this be controlled for?? Furthermore, the finding that individual activity inhibited MVPA could that simply be as a result of the pedagogy that was used to teach those activities?

- We expect that the degree to which various pedagogical processes combine to influence students' MVPA will differ depending on factors such as the lesson context, teacher style, environment, resources, etc. For this reason it would be very challenging to isolate and attempt to control for the effects of specific factors. The SOFIT+ index score is designed to address this issue by providing a composite metric. In response to the comment about Individual Activities, in theory it is true that individual lesson pedagogy could be most influential. However, the pattern in the

data was consistent across the observed lessons even though individual teacher and school variation in pedagogy would be expected.

The authors have referenced Lubans et al SAAFE principles, how does this study dispel or reaffirm their SAAFE principles more broadly? Which should we use at a population level in professional development for teachers?

- We believe that the SOFIT+ methodology is broadly consistent with the SAAFE principles [27] as it can be used to encourage and educate teachers about practices which promote student MVPA. Student engagement in MVPA would ideally occur in a supportive and fun environment that allows for autonomous student participation. The degree to which competition is emphasised (Fair principle) would depend on the nature of the PE activities and how the teachers organise them. In terms of application to teacher professional development SOFIT+ provides guidance on specific teaching practices related to activity promotion, while SAAFE describes a framework of broader guidelines that complement SOFIT+. The SAAFE principles could be useful as an overarching professional development framework with SOFIT+ used as a complementary tool to quantify and evaluate progress on specific elements of SAAFE (i.e., 'Active', 'Fair', Enjoyable' principles).

The limitations are quite brief. Missing details from the methods (see above) may raise additional limitations for the author to consider particularly around representativeness, robustness of associations etc.

- The word limit constraints meant that we could not discuss all of the study limitations in depth. We did though highlight what we felt were the most significant limitations, which included reference to the convenience sample which inhibited the representativeness of the data, and resultant limited generalisability of the findings. Although the number of observed lessons is modest the analyses are rigorous and are based upon over 1,000 scans.

<u>References</u>

1. Welk GJ. Use of accelerometry-based activity monitors to assess physical activity, in *Physical Activity Assessments for Health-Related Research*. Welk GJ, ed^eds. Champaign, IL, Human Kinetics, 2002.

2. Fairclough SJ, Noonan R, Rowlands AV, Van Hees V, Knowles Z, Boddy LM. Wear Compliance and Activity in Children Wearing Wrist- and Hip-Mounted Accelerometers. *Medicine & Science in Sports & Exercise.* 2016; 48(2):245-253.
3. Rowlands AV, Cliff DP, Fairclough SJ, et al. Moving Forward with Backward Compatibility: Translating Wrist Accelerometer Data. *Medicine & Science in Sports & Exercise.* 2016; 48(11):2142-2149.
4. Migueles JH, Cadenas-Sanchez C, Ekelund U, et al. Accelerometer Data Collection and Processing Criteria to Assess Physical Activity and Other Outcomes: A Systematic Review and Practical Considerations. *Sports Medicine.* 2017:1-25.
5. Weaver RG, Webster CA, Erwin H, et al. Modifying the System for Observing Fitness Instruction Time to Measure Teacher Practices Related to Physical Activity Promotion: SOFIT+. *Measurement in Physical Education and Exercise Science.* 2016; 20(2):121-130.
6. Hanggi JM, Phillips LR, Rowlands AV. Validation of the GT3X ActiGraph in children and comparison with the GT1M ActiGraph. *J Sci Med Sport.* 2013; 16(1):40-44.
7. Kim Y, Barry VW, Kang M. Validation of the ActiGraph GT3X and activPAL Accelerometers for the Assessment of Sedentary Behavior. *Measurement in Physical Education and Exercise Science.* 2015; 19(3):125-137.
8. McGarty AM, Penpraze V, Melville CA. Calibration and Cross-Validation of the ActiGraph wGT3X+ Accelerometer for the Estimation of Physical Activity Intensity in Children with Intellectual Disabilities. *PLoS ONE.* 2016; 11(10):e0164928.
9. Evenson KR, Catellier DJ, Gill K, Ondrak KS, McMurray RG. Calibration of two objective measures of physical activity for children. *J Sports Sci.* 2008; 26(14):1557-1565.
10. Trost SG, Loprinzi PD, Moore R, Pfeiffer KA. Comparison of accelerometer cut-points for predicting activity intensity in youth. *Med Sci Sports Exerc.* 2011; 43(7):1360-1368.
11. McKenzie T, Sallis JF, Nader PR. SOFIT - System for Observing Fitness Instruction Time. *Journal of Teaching in Physical Education.* 1991; 11(2):195-205.
12. Ridgers ND, Stratton G, McKenzie T. Reliability and Validity of the system for observing children's activity and relationships during play (SOCARP). *J Phys Act Health.* 2010; 7:17-25.
13. Brown WH, Pfeiffer KA, McIver KL, Dowda M, Almeida MJCA, Pate RR. Assessing Preschool Children's Physical Activity: The Observational System for Recording Physical Activity in Children-Preschool Version. *Research Quarterly for Exercise and Sport.* 2006; 77(2):167-176.
14. McNamee J, Van Der Mars H. Accuracy of momentary time sampling: A comparison of varying interval lengths using SOFIT2005.
15. McKenzie T, Sallis JF, Nader PR, et al. BEACHES: an observational system for assessing children's eating and physical activity behaviors and associated events. *Journal of applied behavior analysis.* 1991; 24(1):141-151.
16. Mcclain J, Abraham T, Brusseau T, Tudor-Locke C. Epoch length and accelerometer outputs in children: comparison to direct observation. *Medicine+ Science in Sports+ Exercise.* 2008; 40(12):2080.
17. Saint-Maurice P, Welk G, Ihmels M, Krapfl J. Validation of the SOPLAY direct observation tool with an accelerometry-based physical activity monitor. *Journal of physical activity & health.* 2011; 8(8):1108-1116.
18. Heath EM, Coleman KJ, Lensegrav TL, Fallon JA. Using momentary time sampling to estimate minutes of physical activity in physical education: validation of scores for the system for observing fitness instruction time. *Research quarterly for exercise and sport.* 2006; 77(1):142-146.
19. Pope RP, Coleman KJ, Gonzalez EC, Barron F, Heath EM. Validity of a revised system for observing fitness instruction time (SOFIT). *Pediatric Exercise Science.* 2002; 14(2):135-146.

20. Rowe P, van Der Mars H, Schuldheisz J, Fox S. Measuring students' physical activity levels: validating SOFIT for use with high-school students. *Journal of Teaching in Physical Education.* 2004; 23(3):235-251.
21. Rowe P, Schuldheisz J, Van der Mars H. Measuring physical activity in physical education: validation of the SOFIT direct observation instrument for use with first to eighth grade students. *Pediatric Exercise Science.* 1997; 9(2):136-149.
22. Scruggs PW. A comparative analysis of pedometry in measuring physical activity of children. *Medicine and science in sports and exercise.* 2007; 39(10):1837-1846.
23. Weaver RG, Beets MW, Webster C, Huberty J. System for Observing Staff Promotion of Activity and Nutrition (SOSPAN). *Journal of physical activity & health.* 2014; 11(1):173 – 185.
24. Bailey RC, Olson J, Pepper SL, Porszaz J, Barstow TJ, Cooper DM. The level and tempo of children's physical activities: An observational study. *Medicine and Science in Sports and Exercise.* 1995; 27:1033-1041.
25. Baquet G, Stratton G, Van Praagh E, Berthoin S. Improving physical activity assessment in prepubertal children with high-frequency accelerometry monitoring: a methodological issue. *Preventive Medicine.* 2007; 44(2):143-147.
26. Vale S, Santos R, Silva P, Soares-Miranda L, Mota J. Preschool children physical activity measurement: importance of epoch length choice. *Pediatr Exerc Sci.* 2009; 21(4):413-420.
27. Lubans DR, Lonsdale C, Cohen K, et al. Framework for the design and delivery of organized physical activity sessions for children and adolescents: rationale and description of the 'SAAFE' teaching principles. *International Journal of Behavioral Nutrition and Physical Activity.* 2017; 14(1):24.

***Title page (including all author details and affiliations)**

Title:	Validation of an observation tool to assess physical activity-promoting physical education lessons in high schools: SOFIT+
Authors:	Stuart J. Fairclough, PhD[a, b] R. Glenn Weaver, PhD[c] Siobhan Johnson, BA[a] Jack Rawlinson, BA[a]
Affiliations	[a] Department of Sport and Physical Activity, Edge Hill University, St Helens Road, Ormskirk, Lancs, UK [b] Department of Physical Education and Sports Science, University of Limerick, Limerick, Ireland [c] Department of Exercise Science, Arnold School of Public Health, University of South Carolina, Columbia, South Carolina
Corresponding Author:	Stuart J. Fairclough (stuart.fairclough@edgehill.ac.uk)
Manuscript word count:	3046 (extra words agreed with editorial office)
Abstract word count:	234
Number of tables:	3
Number of figures:	0

Abstract

Objectives. SOFIT+ is an observation tool to measure teacher practices related to moderate-to-vigorous physical activity (MVPA) promotion during physical education (PE). The objective of the study was to examine the validity of SOFIT+ during high school PE lessons.

Design. This cross-sectional, observational study tested the construct validity of SOFIT+ in boys' and girls' high school PE lessons.

Methods. Twenty-one PE lessons were video-recorded and retrospectively coded using SOFIT+. Students wore hip-mounted accelerometers during lessons as an objective measure of MVPA. Multinomial logistic regression was used to estimate the likelihood of students engaging in MVPA during different teacher practices represented by observed individual codes and a combined SOFIT+ index-score.

Results. Fourteen individual SOFIT+ variables demonstrated a statistically significant relationship with girls' and boys' MVPA. Observed lesson segments identified as high MVPA-promoting were related to an increased likelihood of girls engaging in 5-10 (OR = 2.86 [95%CI 2.41-3.40]), 15-25 (OR = 7.41 [95%CI 6.05-9.06]), and 30-40 (OR = 22.70 [95%CI 16.97-30.37]) seconds of MVPA. For boys, observed high-MVPA promoting segments were related to an increased likelihood of engaging in 5-10 (OR = 1.71 [95%CI 1.45-2.01]), 15-25 (OR = 2.69 [95%CI 2.31-3.13]) and 30-40 (OR = 4.26 [95%CI 3.44-5.29]) seconds of MVPA.

Conclusions. Teacher practices during high school PE lessons are significantly related to students' participation in MVPA. SOFIT+ is a valid and reliable tool to examine relationships between PE teacher practices and student MVPA during PE.

Keywords: Adolescent, Health, Students, MVPA, Validity, Teaching

Introduction

Physical inactivity is observed globally [1] with approximately 20% of adolescents accumulating the recommended 60 minutes/day of moderate-to-vigorous physical activity (MVPA).[2] An age-related decline in objectively measured MVPA is observed starting as young as 6-7 years [3] with activity levels lowest among adolescents.[4] Age-appropriate physical activity opportunities during school are advocated to enhance the health and wellbeing of students [5] through comprehensive school physical activity programmes.[6] Physical education (PE) is key to such programmes, which aims to influence physical activity participation directly through activity engagement during lessons, and indirectly through development of skills, knowledge, and confidence that can enable lifelong activity participation.[7]

Students' MVPA during PE varies, depending on lesson content,[8] student ability,[8] the teaching environment,[9] and grade level.[10] A recent systematic review reported that high school students spent 40.5% of lesson time in MVPA during PE,[10] which is lower than the 50% threshold recommended in the United Kingdom [12] and United States.[11, 12] When MVPA in high school PE is expressed in absolute terms (i.e., minutes), it can however significantly contribute towards adolescents achieving MVPA guidelines.[10] MVPA during PE occurs within a pedagogical context which targets psychomotor, cognitive and affective outcomes. [10] Thus, evaluation of PE quality relative to physical activity promotion requires additional indicators which complement MVPA.

The System for Observing Fitness Instruction Time (SOFIT) [13] is a valid and reliable observation instrument which provides valuable data to inform researchers' and practitioners' understanding of the relationship between PE pedagogy and MVPA engagement during PE. Recently, SOFIT was modified (SOFIT+) to comprehensively measure teacher practices that both promote and

limit MVPA during PE.[14] In addition to Lesson Context, SOFIT+ includes sections to capture Activity Contexts, Teacher Behaviours, and Activity Management (Supplementary file 1). Analysis of 20 US elementary school PE lessons demonstrated that SOFIT+ produced valid and reliable data related to teacher practices for optimising students' MVPA during PE.[14] It is unknown though how applicable SOFIT+ is in high school PE lessons, which can vary from elementary school in respect of structure, activity content, physical environment, and student engagement. The study purpose therefore, was to examine the validity of SOFIT+ during high school PE lessons.

Methods

A detailed description of the SOFIT+ instrument is provided elsewhere.[14] In brief, SOFIT+ is a systematic observation instrument designed to capture teacher practices related to students' engagement in MVPA (Supplementary file 1). SOFIT+ consists of two phases: During phase one *Lesson Context* and *Activity Context* are coded, while *Teacher Behaviours* and *Activity Management* variables are captured in phase two. Phases one and two are completed continuously and consecutively in 20-second intervals throughout the entirety of the observed lesson. At the lesson's end an exit checklist is completed to record lesson location, weather, equipment amount and type, and lesson content.

Ethical approval for this cross-sectional study was granted by Edge Hill University's Sport and Physical Activity Department Research Ethics Committee (SPA-REC-2015-367). Informed consent was collected from eight PE teachers from a convenience sample of two high schools located in medium-high socioeconomic-status and urban areas of West Lancashire, northwest England. Students in school years 7-9 (11-14 years old) were then invited to participate, after

which passive informed assent (students) and consent (parents) were obtained. Twenty-one single-gender PE lessons (i.e., only girls or only boys) taught by the usual teacher were observed over nine school days from January-February 2017.

On observation days, upon arrival to PE class, each student received an ActiGraph GT3x+ accelerometer (ActiGraph; Pensacola, FL) set to record at 30Hz, which was attached to their right hip via an elastic belt. The time that the teacher commenced teaching the lesson was noted by a research assistant. Students then participated in their normally scheduled PE lessons, which were video-recorded using a tripod-mounted GoPro HERO4 camera (GoPro; San Mateo, CA) which captured the entire teaching area. At the end of the lesson, the time was recorded and the accelerometers were collected. SOFIT+ observations were completed via Pendragon Forms Universal© on a Samsung Galaxy Tablet© by viewing the videos. Two lessons were lost due to failed audio recording and accelerometer malfunction, leaving 19 PE lessons in the final sample.

Two trained observers completed all SOFIT+ observations. Training was conducted via Skype (Microsoft; Redmond, WA) by the second author, and consisted of classroom training and video analysis. During the 60 minute classroom training observers reviewed the study protocols, familiarised themselves with the SOFIT+ instrument, committed observational codes to memory, and practiced coding video recorded PE lessons. During video analysis the observers practiced coding SOFIT+ training videos of PE lessons. The criterion for inter-rater agreement was set at ≥80% using interval-by-interval agreement for each category.[15] To establish initial reliability both observers coded SOFIT+ training videos until the criterion percent agreement was met for each instrument code (i.e., eight lessons coded). Both observers were proficient in the use of SOFIT+

after approximately 14 hours of training, after which, data collection in the schools commenced. During the data collection period seven additional lessons, which were included in the sample for this study, were coded by both observers independently, to ensure observer drift did not occur.[15]

Content validity of the SOFIT+ variables (representing teacher practices) has been established previously in elementary schools;[14] this study explored the relationship between the SOFIT+ variables and high school students' accelerometer-derived MVPA. The presence or absence of teacher practices was compared to MVPA using the time-stamped aligned data. One SOFIT+ scan was completed every 40 seconds (i.e., 20 seconds to complete phases one and two, respectively) with data from one accelerometer epoch collected every 5 seconds. This created 40 second segments where one SOFIT+ scan and eight 5 second accelerometer epochs aligned. To test the construct validity of the SOFIT+ instrument in high PE lessons, two methods were used. First, the individual MVPA-promoting practices were examined in relation to students' MVPA. Second, an index-score, representing the total number of MVPA-promoting practices, was created for each 40 second segment. The process to create the index-score described in the original SOFIT+ study [14] is briefly described here: the index-score included the *Lesson Context*, *Activity Context*, and *Teacher Behaviour* variables . However, Fitness, Skill Practice, Game Play, and Free Play were not included as they overlap conceptually with the broader Motor Content *Lesson Context*. Thus, 16 variables were included in the index-score for a possible range of scores of 0 (i.e., all MVPA-limiting variables and no MVPA-promoting variables present) to 16 (i.e., all MVPA-promoting variables and no MVPA-limiting variables present). The creation of the index-score allows assessment of the relationship between overall MVPA supportiveness in each 40 second segment and students' MVPA. Similar to the original SOFIT+ study, the limited number of lessons included in this study did not allow for the

examination of their effect on student MVPA, because variables in the Activity Management section are hypothesised to have an effect on students' MVPA at a lesson level.[14]

Accelerometer data were downloaded using Actilife (v 6.13.3, ActiGraph; Pensacola, FL). Wear time was checked then data were scored using the lesson start and end time filters, with PE lesson MVPA calculated from vertical axis counts.[16] All statistical analyses were completed using STATA (v.14.2, College Station, TX). General descriptive statistics were calculated for SOFIT+ variables and accelerometer derived activity data for students. To examine construct validity, all SOFIT+ variables were collapsed into 40 second segments (i.e., one SOFIT+ scan per segment) with each variable dichotomized as present or not present. MVPA data was then stratified into four categories: no MVPA, 5-10 seconds of MVPA, 15-25 seconds of MVPA, and 30-40 seconds of MVPA. The likelihood of students engaging in each category (i.e., 5-10 through 30-40 seconds of MVPA) when a teacher practice was recorded simultaneously was then estimated using multinomial logistic regression with students engaging in no MVPA as the referent group. The 40 second segments were then divided into two groups: high and low MVPA promotion. High MVPA promotion was defined as an index-score at or above the median score, with low as below the median score. Multinomial logistic regression models estimated the relationship of the index-score to students' engagement in 5-10 , 15-25, and 30-40 seconds of MVPA. All statistical models were run separately for boys and girls and accounted for clustering of scans within children.

Results

A total of 225 students (117 boys) were observed across the 19 PE lessons (10 boys' lessons). Girls and Boys lessons lasted for 34.1 (±2.1) and 38.2 (±4.2) minutes on average, respectively.

Boys accumulated 11.9 (±4.5) and girls 11.0 (±4.5) minutes of MVPA on average. Lesson content included soccer (n=5), rugby (n=2), and handball (n=3) for boys, and fitness circuits (n=3), gymnastics or dance (n=3), athletics (n=1), hockey (n=1), and netball (n=1) for girls. Seven and two lessons were taught outdoors for boys (4.3°C) and girls (5.0°C), respectively.

SOFIT+ interrater reliability is presented in supplementary file 1. A total of 1,079 SOFIT+ scans (502 during girls' lessons) were completed. For girls and boys, the most lesson time was spent in Motor Content (57.2% and 60.0%, respectively), followed by Knowledge Content (27.4% and 20.7%, respectively), and General Content (15.4% and 19.3%, respectively; Table 1). Fitness Activities (31.5%) made up the largest proportion of Motor Content for girls while Game Play was the most prominent for boys (36.3%). For girls, Individual Activities (36.2%) were the most common activity structure, followed by Partner Activities (25.8%). For boys, Large-Sided or Whole-Class activities (36.1%) were observed most often. The most common teacher behaviour was Demonstrate/Instructs (73.8%, 80.1% for girls and boys, respectively), while the most commonly observed *Activity Management* practice was using signals (8.3%, 16.1% for girls and boys, respectively).

TABLE 1

Multinomial logistic regression models relating individual SOFIT+ variables and the SOFIT+ index-score to students' engagement in MVPA are presented in Tables 2 and 3. Fourteen SOFIT+ variables demonstrated a statistically significant relationship with girls' MVPA. Six variables were related to an increased or decreased likelihood of girls engaging in 5-10, 15-25, and 30-40 seconds of MVPA. General Content displayed an increased likelihood for girls to engage in MVPA for 5-10 seconds (OR = 1.29 [95%CI 1.03-1.61]) but a decreased likelihood of engagement in 15-25 (OR = 0.65 [95%CI 0.53-0.81]) and 30-40 seconds (OR = 0.15 [95%CI

0.11-0.21]) of MVPA. An additional six variables demonstrated an increased or decreased likelihood for girls to be engaged in 15-25 and 30-40 seconds of MVPA. Skill Practice was related to an increased likelihood for girls to engage in 5-10 (OR = 3.33 [95%CI 2.52-4.41]) and 15-25 (OR = 4.46 [95%CI 3.18-6.25]) seconds of MVPA. Teacher Off-task was related to a decreased likelihood of girls engaging in 5-10 seconds of MVPA (OR = 0.50 [95%CI 0.25-0.99]). Segments identified as high promoting by the SOFIT+ index were related to an increased likelihood of girls engaging in 5-10 (OR = 2.86 [95%CI 2.41-3.40]), 15-25 (OR = 7.41 [95%CI 6.05-9.06]), and 30-40 (OR = 22.70 [95%CI 16.97-30.37]) seconds of MVPA.

TABLE 2

Fourteen individual SOFIT+ variables demonstrated a significant relationship with boys' MVPA with six variables related to an increased or decreased likelihood for engagement in 5-10, 15-25, and 30-40 seconds of MVPA. Game Play displayed a decreased likelihood for boys to engage in 5-10 seconds of MVPA (OR = 0.78 [95%CI 0.61-0.99]), but an increased likelihood for boys to be engaged in 15-25 (OR = 1.63 [95%CI 1.32-2.02]) and 30-40 (OR = 3.59 [95%CI 2.75-4.68]) seconds of MVPA. An additional seven variables demonstrated an increased or decreased likelihood for boys to be engaged in both 15-25 and 30-40 seconds of MVPA. Segments identified as high promoting by the SOFIT+ index were related to an increased likelihood of boys engaging in 5-10 (OR = 1.71 [95%CI 1.45-2.01]), 15-25 (OR = 2.69 [95%CI 2.31-3.13]) and 30-40 (OR = 4.26 [95%CI 3.44-5.29]) seconds of MVPA.

TABLE 3

Discussion

This study examined the validity of SOFIT+ in high school PE lessons. Where significant associations between MVPA and individual SOFIT+ variables were observed, these were predominantly in the direction hypothesised in the original SOFIT+ study.[14] Motor Content,

Partner Activity, Teacher Promotion of, and Engagement in physical activity were related to an increased likelihood of MVPA engagement, while Knowledge Content, Waiting Activity, students being Off-Task, and teacher Demonstrates/Instructs were associated with reduced likelihood of MVPA.

Skill Practice was positively associated with MVPA but the relationship diminished with longer MVPA bouts. This contrasts with the previous SOFIT+ study which observed inverse but mainly non-significant associations with MVPA.[14] SOFIT+ classifies Skill Practice as MVPA-limiting because the primary goal is skill development, which may involve frequent teacher instruction and feedback,[17] rather than MVPA. In the observed lessons though, skill practice often focused on open skills [18] with emphasis on *active* skills learning, which is effective for enhancing student MVPA during high school PE.[19, 20] Augmented practice episodes may effectively promote skill learning,[21] but a decline in the strength of association between Skill Practice and longer MVPA bouts was observed. Some students in this study lacked sufficient motor skill competence to repeatedly execute skill practices, which may have reduced their physical activity intensity, and contributed to the weaker associations with MVPA bouts > 25 seconds.

During Small-Sided Activities girls were almost twice as likely to be in MVPA than not, while MVPA was less likely for boys during this activity context. Conversely, Large-Sided Activities were positively associated with boys' MVPA and negatively with girls'. These sex-specific differences may be related to delivery of the girls' and boys' lessons, because the girls' Small-Sided Activities were generally co-operative with students working together towards common movement-based goals, which necessitated MVPA (e.g., gymnastics and dance routines, running relays, etc). Moreover, girls' Small-Sided Activities involved Motor Content for 72.2% of

the time, with minimal Waiting Activities and over 25% of time in Game Play. In comparison, during boys' Small-Sided Activities 62.5% of time was devoted to Motor Content, more than half of the small-sided episodes involved Waiting Activity, and Demonstrates/Instructs was the dominant teacher behaviour (83.3%). Large-Sided or Whole Class activities were associated with boys' MVPA engagement which may relate to the lessons being invasion games, which are MVPA-promoting,[22] and 80% of lessons occurred outside where there was maximum space in which to move. Moreover, the outdoor games were soccer and rugby, at which the boys were highly skilled. When students have the motor skill competence and space to participate in invasion games MVPA is likely to be accrued.[22] The girls' Large-Sided and Whole Class Activities were mainly gymnastics and dance performances and fitness and athletics circuit activities, which included numerous inactive periods for recovery, recording performances, etc. It may be that Lesson Contexts, Activity Contexts, and Teacher Behaviours combined to influence students' MVPA during Small-Sided and Large-Sided Activities.

Similar to Small-Sided Activities, divergent associations were observed between Individual Activities and girls' and boys' MVPA (i.e., negative relationship with girls' and positive relationship with boys' MVPA). These relationships may also have been influenced by lesson content. For example, during hockey and fitness circuit lessons, Individual Activities involved locomotor movement, whereby in invasion games, athletics, and dance lessons, Individual Activities were static involving stretching, execution of skills such as throwing, and movement choreography, respectively. Individual Activity was hypothesised to be MVPA-promoting in the elementary school SOFIT+ validation [14] on the basis that students undertake motor activities independently of other students' involvement, and which often result from the direct lead of the teacher (i.e., command teaching style).[23] Our findings however demonstrate that in high school PE lessons Individual Activity may also inhibit MVPA.

242

243 Such mixed findings using SOFIT+ were previously observed [14] and were expected because
244 many events occur simultaneously during lessons,[24] which can influence the associations
245 between individual SOFIT+ variables and MVPA. The SOFIT+ index-score accounted for this
246 *simultaneity* by providing a composite metric, which was significantly related to students' MVPA,
247 particularly among girls. This is consistent with the original SOFIT+ findings,[14] though the
248 strength of the observed associations was greater in the present study. These results confirm
249 that no one individual variable accounts for students' MVPA engagement, and that a
250 combination of best practices applied within specific PE lesson content is required to increase
251 MVPA.

252

253 Teachers' time was mainly spent demonstrating/instructing with teachers verbally promoting in-
254 class or out-of-class physical activity less than 10% of time . This concurs with previous
255 studies,[19, 21, 25] suggesting that MVPA promotion is incidental and perceived to be implicit within
256 the content taught, rather than planned and overtly promoted as normalised teacher behaviour.
257 Enhancing children's appreciation of physical activity, as well as their motivation to engage in
258 MVPA requires emphasis on physical activity education and supportive learning experiences to
259 enhance students' perceived physical activity self-efficacy, competence, and enjoyment.[26, 27] PE
260 teachers are uniquely positioned to influence students in this way but greater professional
261 development and teacher education are warranted to turn this potential into reality.[28]

262

263 Strengths of this study include the objective measurement of physical activity, use of SOFIT+,
264 which encompasses a range of teacher practices related to MVPA engagement, and use of the
265 index-score as a method of examining overall MVPA supportiveness. Limitations included the

positioning of the accelerometers on the hip which meant static upper body and supine leg movements were not detected, and therefore MVPA may have been underestimated in some lessons. Further, we used Evenson et al.'s [16] MVPA cutpoints due to their reported classification accuracy,[29] but recognise that applying alternative cutpoints may have resulted in significantly different MVPA estimates.[30] The generalisability of the results is limited by the modest convenience sample, although the PE content of the lessons were typical of winter activities taught in UK high schools. The small lesson sample size also precluded testing the construct validity of the Activity Management variables.

Conclusion

Teacher practices during high school PE lessons are related to students' participation in MVPA. The direction and magnitude of these associations for individual teacher practices varies depending on PE lesson content, setting, and student engagement. The SOFIT+ index-score provides a valid composite metric by which to examine the relationships between a range of physical activity-promoting and -limiting practices and student MVPA.

Practical implications

- Practitioners can use SOFIT+ as a tool to examine modifiable teacher best-practices related to increasing student MVPA.
- SOFIT+ can be used to measure the effectiveness of PE professional development programmes focused on physical activity promotion.
- SOFIT+ is a valid methodology for quantifying relationships between PE pedagogical practices and student MVPA in observational and intervention studies.

Acknowledgements

The authors thank the teachers and students for their participation in the study. No financial assistance was provided for this study.

References

1. Tremblay MS, Gray CE, Akinroye K, et al. Physical activity of children: A Global Matrix of Grades comparing 15 countries. *J Phys Act Health* 2014; 11(Suppl. 1):S113-S125.
2. Hallal PC, Andersen LB, Bull FC, et al. Global physical activity levels: surveillance progress, pitfalls, and prospects. *Lancet* 2012; 380(9838):247-257
3. Farooq MA, Parkinson KN, Adamson AJ, et al. Timing of the decline in physical activity in childhood and adolescence: Gateshead Millennium Cohort Study. *Br J Sports Med* Online First 2017; doi:10.1136/bjsports-2016-096933
4. Cooper A, Goodman A, Page A, et al. Objectively measured physical activity and sedentary time in youth: the International children's accelerometry database (ICAD). *Int J Behav Nutr Phys Activity* 2015; 12(1):113.
5. Institute of Medicine. 2013 Educating the Student Body: Taking Physical Activity and Physical Education to School. Available at: http://www.iom.edu/~/media/Files/Report Files/2013/Educating-the-Student-Body/EducatingTheStudentBody_Insert.pdf. Accessed 2 December 2016.
6. Centers for Disease Control and Prevention. *Comprehensive School Physical Activity Programs: A Guide for Schools*, Atlanta, GA, US Department of Health and Human Services, 2013.
7. UNESCO. *World-Wide Survey of School Physical Education. Final Report 2013*, Paris, UNESCO, 2014.

8. Fairclough SJ, Stratton G. Physical activity levels in middle and high school physical education: a review. *Pediatr Exerc Sci* 2005; 17:217-236.
9. Fairclough SJ, Stratton G. Effects of a physical education intervention to improve student activity levels. *Phys Educ Sport Pedagogy* 2006; 11(1):29-44.
10. Hollis JL, Sutherland R, Williams AJ, et al. A systematic review and meta-analysis of moderate-to-vigorous physical activity levels in secondary school physical education lessons. *Int J Behav Nutr Phys Activity* 2017; 14(1):52.
11. US Department of Health and Human Services. *Healthy People 2020*, Washington DC: USDHHS, 2010.
12. Association for Physical Education. Health Position Paper. Available at http://www.afpe.org.uk/physical-education/wp-content/uploads/afPE_Health_Position_Paper_Web_Version.pdf. Accessed 2 December 2016.
13. McKenzie TL, Sallis JF, Nader PR. SOFIT: System for Observing Fitness Instruction Time. *J Teach Phys Educ* 1991; 11:195-205.
14. Weaver RG, Webster CA, Erwin H, et al. Modifying the System for Observing Fitness Instruction Time to measure teacher practices related to physical activity promotion: SOFIT+. *Meas Phys Educ Exerc Sci* 2016; 20(2):121-130.
15. Ridgers ND, Stratton G, McKenzie TL. Reliability and validity of the System for Observing Children's Activity and Relationships During Play (SOCARP). *J Phys Activity Health* 2010; 7(1):17-25.
16. Evenson KR, Catellier DJ, Gill K, et al. Calibration of two objective measures of physical activity for children. *J Sports Sci* 2008; 26(14):1557-1565.
17. McKenzie TL. SOFIT. System for Observing Fitness Instruction Time. Description and procedures manual. Available at:

http://activelivingresearch.org/sites/default/files/SOFIT_Protocols_05.01.15.pdf. Accessed 2 December 2016.

18. Wang C-H, Chang C-C, Liang Y-M, et al. Open vs. closed skill sports and the modulation of inhibitory control. *PLOS ONE* 2013; 8(2):e55773.

19. Fairclough SJ, Stratton G. Improving health-enhancing physical activity in girls' physical education. *Health Educ Res* 2005; 20(4):448-457.

20. McKenzie TL, Catellier DJ, Conway TL, et al. Girls' activity levels and lesson contexts in middle school PE: TAAG baseline. *Med Sci Sports Exerc* 2006; 38:1229-1235.

21. McKenzie TL, Marshall SJ, Sallis JF, Conway TL. Student activity levels, lesson context, and teacher behavior during middle school physical education. *Res Q Exerc Sport.* 2000 71(3):249-259.

22. Fairclough SJ, Stratton G. 'Physical education makes you fit and healthy'. Physical education's contribution to young people's physical activity levels. *Health Educ Res* 2005; 20(1):14-23.

23. Mosston M, Ashworth S. *Teaching Physical Education*, New York, Macmillan, 1994.

24. Doyle W. *Classroom Management*, West Lafayette, IN, Kappa Delta Pi, 1980.

25. Mersh R, Fairclough SJ. Physical activity, lesson context and teacher behaviours within the revised English National Curriculum for Physical Education: A case study of one school. *Eur Phys Educ Rev* 2010; 16(1):29-45.

26. Wallhead T, Buckworth J. The role of physical education in the promotion of youth physical activity. *Quest* 2004; 56:285-301.

27. Lubans DR, Lonsdale C, Cohen K, et al. Framework for the design and delivery of organized physical activity sessions for children and adolescents: rationale and description of the ‘SAAFE’ teaching principles. *Int J Behav Nutr Phys Act* 2017; 14(1):24.

28. Weaver RG, Webster CW, Egan CA, Campos CMC, Michael RD, Vazou S. Partnerships for Active Children in Elementary Schools: Outcomes of a two-year pilot study to

increase physical activity during the school day. *Am J Health Promot* In press; doi: 10.1177/0890117117707289.

29. Trost SG, Loprinzi PD, Moore R, Pfeiffer KA. Comparison of accelerometer cut-points for predicting activity intensity in youth. *Med Sci Sports Exerc* 2011; 43(7):1360-1368.

30. Kim Y, Beets MW, Welk GJ. Everything you wanted to know about selecting the "right" Actigraph accelerometer cut-points for youth, but…: A systematic review. *J Sci Med Sport* 2012; 15(4):311-321.

Table 1

Table 1. Incidence of behaviours, management of the physical activity environment, and scheduled activity across total scans

	Girls PE Lessons				Boys PE Lessons			
Variable	Percent of Scans During the PE Class	SD	Percent of Scans During Motor Content	SD	Percent of Scans During the PE Class	SD	Percent of Scans During Motor Content	SD
Lesson context								
General Content	15.4	36.1			19.3	39.5		
Knowledge Content	27.4	44.6			20.7	40.6		
Motor Content	57.2	49.5			60.0	49.0		
Fitness	31.5	46.5	55.2	49.8	8.4	27.7	14.0	34.7
Skill Practice	15.3	36.0	26.7	44.3	17.3	37.9	28.9	45.4
Game play	10.4	30.6	18.2	38.6	36.3	48.1	60.5	49.0
Free play	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Activity context								
Individual activity	36.2	48.1	23.6	42.5	26.5	44.2	12.3	32.9
Partner activity	25.8	43.8	25.2	43.5	4.2	20.0	4.7	21.1
Small sided activity	13.7	34.4	17.3	37.9	33.5	47.2	34.9	47.7
Large sided or whole class activity	24.3	42.9	33.9	47.4	36.1	48.1	48.2	50.0
Waiting Activity	22.0	41.5	25.2	43.5	43.4	49.6	46.5	50.0
Elimination Activity	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Students off task	0.5	7.2	0.6	7.8	0.4	6.3	0.3	5.8
Teacher Behaviours								
Demonstrate/Instructs	73.8	44.0	68.5	46.5	80.1	40.0	73.4	44.2
Promotes Physical Activity	7.5	26.3	10.6	30.8	7.4	26.2	11.6	32.1
PA as Punishment	0.0	0.0	0.0	0.0	0.2	4.5	0.3	5.8
Withholding PA	0.0	0.0	0.0	0.0	0.6	7.7	0.7	8.1
PA Engaged with Students	0.0	0.0	0.0	0.0	4.4	20.5	6.0	23.8
Teacher Off/Other Task	1.7	13.1	3.0	17.2	4.2	20.0	5.0	21.8
Activity Management								
Using signals	8.3	27.6	7.9	27.0	16.1	36.8	19.6	39.8
Retrieving equipment many access points	0.0	0.0	0.0	0.0	0.2	4.5	0.0	0.0
Retrieving equipment one access point	1.0	10.2	0.3	5.5	1.4	11.7	0.7	8.1
Grouping	2.4	15.4	1.2	11.0	4.2	20.0	1.3	11.5
Addressing interruptions public	3.5	18.3	0.9	9.5	4.4	20.5	2.3	15.1
Addressing interruptions private	4.0	19.6	4.5	20.9	1.2	10.9	1.3	11.5

Based upon SOFIT+ 1,079 scans (502 during girls' PE lessons) during 19 PE lessons (10 girls' lessons), 631 scans occurring during motor content (330 during motor content in girls' PE lessons)

Table 2

Table 2. Construct Validity of SOFIT+ for Girls

	5-10 seconds of MVPA	95% CI		15-25 seconds of MVPA	95% CI		30-40 seconds of MVPA	95% CI	
Lesson Context									
General Content -	**1.29**	**(1.03,**	**1.61)**	**0.65**	**(0.53,**	**0.81)**	**0.15**	**(0.11,**	**0.21)**
Knowledge Content -	**0.33**	**(0.28,**	**0.39)**	**0.15**	**(0.12,**	**0.18)**	**0.04**	**(0.03,**	**0.06)**
Motor Content +	**2.35**	**(1.99,**	**2.77)**	**5.99**	**(4.90,**	**7.32)**	**25.41**	**(19.20,**	**33.63)**
Fitness +	1.03	(0.79,	1.35)	**1.81**	**(1.28,**	**2.58)**	**9.36**	**(6.25,**	**14.03)**
Skill Practice -	**3.33**	**(2.52,**	**4.41)**	**4.46**	**(3.18,**	**6.25)**	1.42	(0.92,	2.20)
Game play +	1.43	(0.96,	2.13)	**2.34**	**(1.70,**	**3.22)**	**2.07**	**(1.32,**	**3.24)**
Free play -	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Activity Context									
Individual activity +	0.91	(0.76,	1.08)	**0.52**	**(0.39,**	**0.69)**	**0.24**	**(0.17,**	**0.35)**
Partner activity +	1.04	(0.86,	1.26)	**1.80**	**(1.40,**	**2.31)**	**3.59**	**(2.50,**	**5.14)**
Small sided activity +	**1.72**	**(1.38,**	**2.14)**	**1.94**	**(1.54,**	**2.46)**	**1.79**	**(1.39,**	**2.32)**
Large sided or whole class activity -	**0.54**	**(0.41,**	**0.70)**	**0.51**	**(0.36,**	**0.71)**	**0.45**	**(0.31,**	**0.66)**
Waiting Activity -	**0.62**	**(0.51,**	**0.77)**	**0.36**	**(0.28,**	**0.47)**	**0.17**	**(0.11,**	**0.26)**
Elimination Activity -	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Students Off task -	0.66	(0.36,	1.22)	**0.19**	**0.06**	**0.67**	**0.07**	**(0.01,**	**0.50)**
Teacher Behaviours									
Demonstrates or Instructs -	1.05	(0.89,	1.24)	0.93	(0.80,	1.08)	0.97	(0.83,	1.12)
Promotes Physical Activity +	1.05	(0.75,	1.48)	**1.92**	**(1.39,**	**2.65)**	**2.17**	**(1.60,**	**2.93)**
Physical Activity as Punishment -	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Withholding Physical Activity as Punishment -	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Physical Activity Engaged with Students +	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Teacher Other Task -	**0.50**	**(0.25,**	**0.99)**	0.79	(0.46,	1.34)	1.39	(0.82,	2.37)
Sofit+ Index	**2.86**	**(2.41,**	**3.40)**	**7.41**	**(6.05,**	**9.06)**	**22.70**	**(16.97,**	**30.37)**

"N/A" behavior never observed, thus no model estimated
"-"variable theorized to detract from student's engagement in physical activity and reverse coded in SOFIT+ index
"+"variable theorized to increase student's engagement in physical activity
No epochs in MVPA is the reference group

Table 3

Table 3. Construct Validity of SOFIT+ for Boys

	5-10 seconds of MVPA	95% CI		15-25 seconds of MVPA	95% CI		30-40 seconds of MVPA	95% CI	
Lesson Context									
General Content -	0.87	(0.74,	1.03)	**0.45**	**(0.39,**	**0.52)**	**0.26**	**(0.21,**	**0.32)**
Knowledge Content -	**0.58**	**(0.49,**	**0.68)**	**0.44**	**(0.37,**	**0.53)**	**0.30**	**(0.23,**	**0.39)**
Motor Content +	**1.71**	**(1.46,**	**2.01)**	**3.16**	**(2.65,**	**3.76)**	**5.67**	**(4.34,**	**7.40)**
Fitness +	**1.88**	**(1.38,**	**2.56)**	**2.39**	**(1.70,**	**3.37)**	**2.03**	**(1.45,**	**2.85)**
Skill Practice -	**2.85**	**(2.29,**	**3.55)**	**2.51**	**(2.04,**	**3.07)**	**1.71**	**(1.30,**	**2.26)**
Game play +	**0.78**	**(0.61,**	**0.99)**	**1.63**	**(1.32,**	**2.02)**	**3.59**	**(2.75,**	**4.68)**
Free play -	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Activity Context									
Individual activity +	0.93	(0.78,	1.10)	**0.72**	**(0.59,**	**0.88)**	**0.57**	**(0.43,**	**0.77)**
Partner activity +	0.73	(0.42,	1.28)	**3.05**	**(2.09,**	**4.46)**	**2.58**	**(1.64,**	**4.04)**
Small-sided activity +	0.95	(0.80,	1.13)	**0.43**	**(0.36,**	**0.52)**	**0.25**	**(0.19,**	**0.34)**
Large-sided or whole class activity -	1.31	(0.95,	1.80)	**3.15**	**(2.27,**	**4.39)**	**5.89**	**(3.96,**	**8.75)**
Waiting Activity -	0.88	(0.74,	1.05)	**0.61**	**(0.53,**	**0.72)**	**0.43**	**(0.35,**	**0.52)**
Elimination Activity -	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Students Off task -	0.96	(0.26,	3.55)	1.21	(0.32,	4.57)	0.79	(0.20,	3.06)
Teacher Behaviours									
Demonstrates or Instructs -	**0.69**	**(0.54,**	**0.89)**	**0.46**	**(0.36,**	**0.60)**	**0.35**	**(0.27,**	**0.45)**
Promotes Physical Activity +	0.80	(0.57,	1.13)	1.25	(0.94,	1.67)	**2.16**	**(1.62,**	**2.86)**
Physical Activity as Punishment -	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Withholding Physical Activity as Punishment -	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Physical Activity Engaged with Students +	**2.80**	**(1.78,**	**4.41)**	**2.88**	**(1.81,**	**4.57)**	**1.86**	**(1.05,**	**3.29)**
Teacher Other Task -	1.15	(0.84,	1.55)	**1.52**	**(1.09,**	**2.11)**	**1.59**	**(1.10,**	**2.31)**
Sofit+ Index	**1.71**	**(1.45,**	**2.01)**	**2.69**	**(2.31,**	**3.13)**	**4.26**	**(3.44,**	**5.29)**

1 "N/A" behavior never observed, thus no model estimated
2 "-"variable theorized to detract from student's engagement in physical activity and reverse coded in SOFIT+ index
3 "+"variable theorized to increase student's engagement in physical activity
4 No epochs in MVPA is the reference group

Supplementary Material

Supplementary file 1. SOFIT+ variables and inter-observe agreement

Variable	Operational Definition	Percent Agreement
Lesson context		
General Content -	Students are not intended to be involved in physical education content, including transition, management, and break times[18]	89.7
Knowledge Content -	Primary focus of the lesson is on student acquisition of knowledge related to physical education, not activity engagement[18]	87.3
Motor Content +	Lesson time when the primary focus is on student motor engagement[18]	90.0
Fitness +	Activities whose major purpose is to alter the physical state of the student in terms of cardiovascular endurance, strength, or flexibility[18]	97.6
Skill Practice -	Time devoted to practice of skills[18]	96.9
Game play +	Time devoted to the application of skills in a game or competitive setting[18]	95.7
Free play -	Free choice activity time during which physical education instruction is not intended[18]	99.8
Activity context		
Individual activity +	Students participate in an activity alone, without interacting with others to accomplish a task	89.3
Partner activity +	Students work with partners/groups of two to engage in task	99.3
Small sided activity +	Students are divided into several small groups (no more than 5 per activity or game) instead of one large group.	89.7
Large sided or whole class activity -	Students participate in an activity as a large group, and must interact with others to accomplish specific tasks	99.8
Waiting Activity -	Physical activity provided by the teacher requires that students either 1) must wait for their turn to play/participate or 2) are waiting for the teacher to provide instruction	88.3
Elimination Activity -	The activity eliminates students from physical activity opportunities as it progresses	99.0
Students off task -	One or more students are not engaged in the activity presented by teacher[18] and are disrupting other students	99.0
Teacher Behaviors		
Demonstrate/Instructs -	Models a physical activity for the purpose of student understanding or lectures, describes, prompts, or provides feedback to students related to all physical education content[18]	81.8
Promotes Physical Activity +	Teacher verbally promotes in-class or out of class physical activity, motor skills, or fitness[18]	93.5
Physical Activity as Punishment -	The teacher prescribes physical activity as a punishment for misbehavior[18]	N/A

Withholding Physical Activity -	The teacher removes a student from physical activity or threatens to remove a student from physical activity (present or future) as a consequence for behavior	N/A
Physical Activity Engaged with Students +	The teacher is participating in physical activity with the students	99.7
Teacher Off/Other Task -	The teacher attends to events related to their duties or not related to their responsibilities to the class at hand and is not observing the class behavior[18]	98.8
Activity Management		
Freezing	Teacher provides a signal indicating students should stop the activity and demonstrate they are ready to listen/observe	92.1
Retrieving equipment many access points	Students are moving to the equipment, retrieving a piece of equipment, or moving back to the teaching area in preparation to use the equipment and the teacher has provided more than one access point for equipment	N/A
Retrieving equipment one access point	Students are moving to the equipment, retrieving a piece of equipment, or moving back to the teaching area in preparation to use the equipment and the teacher has provided one access point for equipment	99.5
Grouping	Students are being moved into a different formation	97.9
Addressing interruptions public	The teacher is addressing an interruption (for example: misbehavior, injury) while the entire class is stopped	94.0
Addressing interruptions private	The teacher is addressing an interruption (for example: misbehavior, injury) but the class continues as normal	97.3

***Acknowledgements**

Acknowledgements

The authors thank the teachers and students for their participation in the study. No financial assistance was provided for this study.