

## PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (<http://bmjopen.bmj.com/site/about/resources/checklist.pdf>) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

### ARTICLE DETAILS

<b>TITLE (PROVISIONAL)</b>	A school-based physical activity intervention for older adolescents: Rationale and study protocol for the Burn 2 Learn cluster randomised controlled trial
<b>AUTHORS</b>	Leahy, Angus; Eather, Narelle; Smith, Jordan; Hillman, Charles; Morgan, Philip; Nilsson, Michael; Lonsdale, Chris; Plotnikoff, Ronald C.; Noetel, Michael; Holliday, Elizabeth; Shigeta, Tatsuya; Costigan, Sarah; Walker, F; Young, Sarah; Valkenborghs, Sarah; Gyawali, Prajwal; Harris, Nigel; Kennedy, Sarah; Lubans, David

### VERSION 1 - REVIEW

<b>REVIEWER</b>	Stephen Malden 1. University of Strathclyde, UK 2. University of Edinburgh, UK
<b>REVIEW RETURNED</b>	11-Oct-2018

<b>GENERAL COMMENTS</b>	<p>This paper presents the rationale and protocol for a cluster randomised controlled trial to test the impact of school-based Burn 2 Learn physical activity intervention in adolescents.</p> <p>Overall the manuscript is well written, and the study has been appropriately designed to measure the outcomes of interest.</p> <p>The manuscript would benefit from a small number of minor revisions that would mainly serve to provide the wider readership with more information regarding the nature of physical activity levels in later adolescence, specific literature relating to the rationale for the study, and methodological procedures during data collection; detailed below:</p> <p>Introduction: Do physical activity levels significantly differ by gender in Australian youth? If so, this may be an important point to make in the introduction. Even just as a citation rather than an in depth discussion due to word count constraints.</p> <p>Theoretical framework for the intervention: Line 194- Is there a reference available which demonstrates that secondary schools in NSW do not provide mandatory PA lessons to senior students? For example a national/regional school curriculum document.</p>
-------------------------	--

	<p>Secondary outcomes: Line 252- Please provide a reference for the existing categories of physical activity intensity mentioned here.</p> <p>Body composition- is there any protocol in place to reduce the risk of measurement bias during anthropometric measures? (i.e. repeated measurement?) If so please state this in this section.</p> <p>Statistical analysis: Line 424- Again, as you plan to do a subgroup analysis by sex, the paper may benefit from a brief discussion in the introduction regarding the disparity observed between the decline in PA through adolescence between boys and girls (if the same trend applies to Australian youth). This has been an important focus for a while in adolescent health research so might be good to briefly acknowledge it. Lines 426 to 427- the formatting error makes it difficult to decipher what is being said in this sentence and the start of the following paragraph.</p>
--	---

<b>REVIEWER</b>	Kate Ridley Flinders University Adelaide, Australia
<b>REVIEW RETURNED</b>	31-Oct-2018

<b>GENERAL COMMENTS</b>	<p>This manuscript describes the protocol for a RCT designed to improve upper secondary students' physical, cognitive, and mental health using a high-intensity interval training (HIIT) method administered by teachers in school curriculum time. This study builds on a previously conducted and published pilot trial by aiming to assess the potential of the program to be 'up-scaled' to a larger population. Many physical activity interventions efficacy trials conducted in schools are successful in the short term, but less is known about sustainability and optimal methods for successful implementation at scale. Therefore, a strength of the protocol is the focus on implementation via process evaluations for both teachers (intervention administrators) and students (participants). The study demonstrates innovation in its measures. While physical fitness is the primary outcome, the study employs a number of objective measures to study secondary outcomes such as cognitive and mental health. For example, brain structure/function will be assessed via magnetic resonance imaging and stress will be measured by cortisol concentrations from hair samples. Hence, there is potential for the data collected in this study to contribute to our understanding of both: health associations of vigorous physical activity in children; and intervention design/behaviour change.</p> <p>The manuscript provides a solid rationale for the study and presents a very thorough, detailed, and predominantly clear explanation of the protocol. I only have a few questions and comments that require attention.</p> <p>Abstract: L11 - Please include the full name of the B2L intervention before referring to the abbreviation. L16 - Please add age range of the participants. L31 – It would be helpful to note the health outcomes measured by the cortisol concentrations and multi-model magnetic resonance</p>
-------------------------	---

imaging for readers of the abstract who are unfamiliar with the role of these tests.

#### Introduction

L16 – should the text in the brackets read when healthy adolescents participated.....?

L65 – note that the effect for psychological distress was (presumably) negatively associated or a favourable health effect.

#### Methods

L 98 – The selection criteria state that Grade 11-12 students are eligible for participation. However, the remainder of the methods refer to only grade 11 at baseline. Are only Grade 11 students being recruited at baseline allowing for 12-month follow up at Grade 12?

L 111 – I have a general query about selection bias regarding both teacher and student recruitment. Will the characteristics of teachers and the nature of their classes/students be investigated to examine a possible selection bias? E.g. Teachers with physical education training and students who elect to study an optional Physical Education course may be more conducive open to facilitating and participating in HIIT activities? This could have impact on reach (reaching the most inactive students for instance), recruitment rates, compliance/adherence and your future conclusions regarding success of implementation/scalability.

L163 – The professional learning workshop for teachers seems to be omitted from the Intervention delivery, components and implementation strategies paragraph. Aside from Table 1, I believe it is first mentioned at line 402 under Process Evaluation. Could it be introduced at line 163 when discussing Phase 1?

Order of battery of tests – Participants are being asked to participate in a large number of tests include physical measures and self-report questionnaires. Can you please briefly discuss timing and ordering of the tests? Will all testing occur on the same day? I note the cognitive tests will be undertaken before the fitness assessments, but will the cognitive control tests be routinely administered prior to the other self-reports, or order randomised?

L440 – Please comment on your rationale for using an age-predicted HR max rather than an individualised HR max. What equation will be used to predict HR max? "220 - age", or "208 - (0.7 x age)", or other? As participants will be of the same, or similar age, each individual will be aiming for the same/similar HR during the HIIT sessions. Was this uniformity considered in light of your "Fair" component of the SAAFE teaching principles, i.e. "encourage students to modify exercises to personal fitness and ability level" & "encourage self- rather than peer-comparison"? As the participants are undertaking a maximal test (shuttle run test) prior to the intervention individualised HR max could have been collected via HR telemetry during this test. Was this procedure considered?

#### Discussion

Line 468 – Given the length of the manuscript you might like to reconsider re-stating the Beets theory here as it is already discussed in the theoretical framework at line 189.

References – Some references are missing journal volume and page number details, e.g. ref #s 9, 72 for instance.

	Tables – Please add table legends defining abbreviations, such as B2L and SAAFE, so the tables can stand-alone.
--	---

<b>REVIEWER</b>	Dan Benardot, PhD, RD, LD, FACSM *Georgia State University, USA *Emory University, USA
<b>REVIEW RETURNED</b>	21-Dec-2018

<b>GENERAL COMMENTS</b>	<p>Abstract: First use of abbreviation (B2L) not defined. HIIT, which is a well-known abbreviation, is defined in the first sentence.</p> <p>Page 6: Cortisol may be elevated when subjects are in a low energy balance state (see Torstveit et al. 2018; Fahrenholtz et al. 2018; and Mountjoy et al. [RED-S] 2018), yet this protocol includes no strategy for assessing the state of energy balance (i.e., are they in a state of Relative Energy Deficiency in Sport [RED-S]?) at the time of the HIIT session. Including such an assessment strategy would help to better understand the potential benefits of HIIT in subjects who are in or not in a good energy balanced state. Performing HIIT when in a low energy balance state is likely to negatively influence the potential benefits of the exercise. Since subjects are from different schools, the HIIT may be performed at different times that influence energy balance. For instance, subjects performing HIIT shortly after lunch are likely to be in a better energy balance and achieve better outcomes, than those who perform HIIT just before lunch.</p> <p>Page 8: Delete the word "overall" as it is ill-defined and diminishes the power of the sentence.</p> <p>Page 12: This appears to be an appropriate protocol for randomization of Cohort 1 and Cohort 2 schools.</p> <p>Page 15: The body composition protocol fails to adequately assess the relative change in body fat/lean mass as a result of the protocol. BMI has been shown to be a poor measure for assessing relative body fatness in children. Suggest adding a validated strategy for assessing body composition.</p> <p>Page 18: There is a nutritional component to psychological well-being that is not addressed. The brain is highly reliant on blood glucose for energy, and blood glucose can drop rapidly during the adolescent growth spurt and can also drop rapidly with physical activity if not actively supported with an appropriate nutritional protocol (eg. sipping on a sports beverage occasionally during activity, etc.).</p> <p>Page 20: There is a formatting issue that makes the sentence unclear.</p> <p>General: There is a strong relationship between nutritional state and the potential benefits that can be achieved from physical activity. Failure to account for this relationship may diminish the usefulness of the results.</p>
-------------------------	---

## VERSION 1 – AUTHOR RESPONSE

Reviewer: 1

1.1: This paper presents the rationale and protocol for a cluster randomised controlled trial to test the impact of school-based Burn 2 Learn physical activity intervention in adolescents. Overall the manuscript is well written, and the study has been appropriately designed to measure the outcomes of interest.

We thank the Reviewer for their positive comments and taking the time to review our manuscript.

1.2: The manuscript would benefit from a small number of minor revisions that would mainly serve to provide the wider readership with more information regarding the nature of physical activity levels in later adolescence, specific literature relating to the rationale for the study, and methodological procedures during data collection; detailed below:

We thank the Reviewer for their insight provided and have addressed their comments in the sections below.

1.3: Do physical activity levels significantly differ by gender in Australian youth? If so, this may be an important point to make in the introduction. Even just as a citation rather than an in depth discussion due to word count constraints.

In response to the Reviewer's suggestion, we have added the following citation describing the differences between males and females in relation to meeting physical activity recommendations.

"Recent data gathered from Australian secondary schools indicates that adolescent girls are less likely to meet physical activity recommendations than boys. 4"

1.4: Line 194- Is there a reference available which demonstrates that secondary schools in NSW do not provide mandatory PA lessons to senior students? For example a national/regional school curriculum document.

We have added the following link and citation that demonstrates NSW government schools do not provide mandatory physical activity for senior school students (i.e., Grades 11 and 12):

<https://education.nsw.gov.au/policy-library/policies/sport-and-physical-activity-policy>

1.5: Line 252- Please provide a reference for the existing categories of physical activity intensity mentioned here.

We now cite Chandler et al's study to describe the physical activity intensity thresholds used.

Chandler, J., Brazendale, K., Beets, M., & Mealing, B. (2016). Classification of physical activity intensities using a wrist-worn accelerometer in 8–12-year-old children. *Pediatric obesity*, 11(2), 120-127.

1.6: Body composition- is there any protocol in place to reduce the risk of measurement bias during anthropometric measures? (i.e. repeated measurement?) If so please state this in this section.

We thank the Reviewer for their suggestion. The following has been added to the manuscript:

"Both weight and height will be measured twice to reduce the risk of measurement error. A third measurement will occur should there be a difference of >0.1 kg for weight, and >0.3 cm for height between the first and second measurement."

1.7: Line 424- Again, as you plan to do a subgroup analysis by sex, the paper may benefit from a brief discussion in the introduction regarding the disparity observed between the decline in PA through adolescence between boys and girls (if the same trend applies to Australian youth). This has been an important focus for a while in adolescent health research so might be good to briefly acknowledge it.

This a valid point raiser by the reviewer, however, our intervention targets both boys and girls. Due to our already lengthy introduction and the BMJ Open word limit restrictions, it is not possible to add this information to the manuscript.

1.8: Lines 426 to 427- the formatting error makes it difficult to decipher what is being said in this sentence and the start of the following paragraph.

We thank the Reviewer for bringing this to our attention. The text should read as follows:

“Subgroup analyses will be conducted for the following variables if the significance of the group-by-moderator interaction is  $\leq .10$ : Socio-economic status (low, medium, high), sex (male, female), baseline weight status (not overweight, overweight/obese), baseline psychological distress (using established cut-offs from the Strengths and Difficulties Questionnaire), and baseline CRF (using FITNESSGRAM fitness standards). Compared to complete case analyses, mixed models include available data for all participants in the analysis and are thus both more efficient and robust to bias. Mixed model analyses are consistent with the intention-to-treat principle, assuming the data are missing at random 5.”

Reviewer: 2

2.1: This manuscript describes the protocol for a RCT designed to improve upper secondary students' physical, cognitive, and mental health using a high-intensity interval training (HIIT) method administered by teachers in school curriculum time. This study builds on a previously conducted and published pilot trial by aiming to assess the potential of the program to be 'up-scaled' to a larger population. Many physical activity intervention efficacy trials conducted in schools are successful in the short term, but less is known about sustainability and optimal methods for successful implementation at scale. Therefore, a strength of the protocol is the focus on implementation via process evaluations for both teachers (intervention administrators) and students (participants). The study demonstrates innovation in its measures. While physical fitness is the primary outcome, the study employs a number of objective measures to study secondary outcomes such as cognitive and mental health. For example, brain structure/function will be assessed via magnetic resonance imaging and stress will be measured by cortisol concentrations from hair samples. Hence, there is potential for the data collected in this study to contribute to our understanding of both: health associations of vigorous physical activity in children; and intervention design/behaviour change.

We thank the Reviewer for taking their time to review our manuscript.

2.2: The manuscript provides a solid rationale for the study and presents a very thorough, detailed, and predominantly clear explanation of the protocol. I only have a few questions and comments that require attention.

We thank the Reviewer for their positive comments, which we have addressed below:

2.3: Abstract. L11 - Please include the full name of the B2L intervention before referring to the abbreviation.

This has been amended and now reads:

“Methods and analysis: The Burn 2 Learn (B2L) intervention will be evaluated using a two-arm parallel group cluster randomised controlled trial with allocation occurring at the school level (to treatment or wait-list control).”

2.4: L16 - Please add age range of the participants.

The age of participants has been added. We now state:

“The trial will aim to recruit ~720 senior school students (aged 16-18 years) from 20 secondary schools (i.e., 10 schools per cohort).”

2.5: L31 – It would be helpful to note the health outcomes measured by the cortisol concentrations and multi-modal magnetic resonance imaging for readers of the abstract who are unfamiliar with the role of these tests.

This section has been amended and now reads:

“A subsample of students will (i) provide hair samples to determine their accumulated exposure to stressful events and (ii) undergo multi-modal magnetic resonance imaging to examine brain structure and function.”

2.6: Introduction. L16 – should the text in the brackets read <u>when</u> healthy adolescents participated.....?

The text has been amended to the following:

For example, a recent longitudinal study involving a large sample of adolescents found that cardio-metabolic risk declined in a dose-response manner with increasing vigorous physical activity in adolescents (healthy adolescents accumulated at least 7mins of vigorous activity each day), but not with increased volume of light or moderate physical activity.

2.7: L65 – note that the effect for psychological distress was (presumably) negatively associated or a favourable health effect.

The text has been amended and reads as follows:

“Favourable intervention effects were found for CRF, lower-body muscular power (increases) and psychological distress (decrease) in the hypothesised directions”.

2.8: Methods. L 98 – The selection criteria state that Grade 11-12 students are eligible for participation. However, the remainder of the methods refer to only grade 11 at baseline. Are only Grade 11 students being recruited at baseline allowing for 12-month follow up at Grade 12?

Only Grade 11 students will be recruited for the study at baseline, which will allow for a 12-month follow up during Grade 12. In our methods section we state:

Eligible participants will be Grade 11 students who are taught by one of the B2L school champions.

2.9: L 111 – I have a general query about selection bias regarding both teacher and student recruitment. Will the characteristics of teachers and the nature of their classes/students be investigated to examine a possible selection bias? E.g. Teachers with physical education training and students who elect to study an optional Physical Education course may be more conducive open to facilitating and participating in HIIT activities? This could have impact on reach (reaching the most inactive students for instance), recruitment rates, compliance/adherence and your future conclusions regarding success of implementation/scalability.

We thank the Reviewer for their valid point regarding potential reach of the program. The teaching discipline of B2L school champions (i.e., Mathematics, English, Health and Physical Education) and class list of the participating students will be recorded, examined and presented in subsequent publications. We do acknowledge that teachers and students with an interest in Physical Education may be more likely to engage in the program. However, this is likely to be the case for most schools, and would therefore be representative of a normal secondary school context.

2.10: L163 – The professional learning workshop for teachers seems to be omitted from the Intervention delivery, components and implementation strategies paragraph. Aside from Table 1, I believe it is first mentioned at line 402 under Process Evaluation. Could it be introduced at line 163 when discussing Phase 1?

The following has been added in the 'Intervention delivery, components and implementation strategies' paragraph:

“During Phase 1, school champions will attend a one day professional learning workshop led by the research team. The workshop will focus on providing the school champions with the knowledge and skills needed to introduce students to HIIT and develop their competency.”

2.11: Order of battery of tests – Participants are being asked to participate in a large number of tests include physical measures and self-report questionnaires. Can you please briefly discuss timing and ordering of the tests? Will all testing occur on the same day? I note the cognitive tests will be undertaken before the fitness assessments, but will the cognitive control tests be routinely administered prior to the other self-reports, or order randomised?

All measures will be conducted at the study schools on the same day except for the multi-modal magnetic resonance imaging, which will only be measured in a subset of participants. Due to the limited equipment available (i.e., laptops/electronic tablets) and to minimise burden on normal school practice, the self-report and cognitive measures will be completed in a randomised order. An approximate timing for completion of the testing battery has also been added. We now state:

Measures and data collection

“Apart from multi-modal magnetic resonance imaging, all assessments will be conducted at the study schools on the same day by trained research assistants, who will be blinded to group allocation at all time-points. Demographic information and self-report measures will be collected using electronic tablets under exam-like conditions. Cognitive testing will occur on University laptops and participants will receive instructions and practice prior to performing each of the cognitive tasks. Self-report and cognitive measures will occur prior to fitness assessments in a randomised order. Anthropometric assessments will be conducted in a sensitive manner by same-sex research staff when possible. Research assistants will provide a brief verbal description and demonstration of each fitness test prior to commencement. The timing for participants to complete all testing is approximately 90 minutes.”

2.12: L440 – Please comment on your rationale for using an age-predicted HR max rather than an individualised HR max. What equation will be used to predict HR max? "220 - age", or "208 - (0.7 x age)", or other? As participants will be of the same, or similar age, each individual will be aiming for the same/similar HR during the HIIT sessions. Was this uniformity considered in light of your "Fair" component of the SAAFE teaching principles, i.e. "encourage students to modify exercises to personal fitness and ability level" & "encourage self- rather than peer-comparison"? As the participants are undertaking a maximal test (shuttle run test) prior to the intervention individualised HR max could have been collected via HR telemetry during this test. Was this procedure considered?

We decided to use an age-predicted heart rate maximum formula (220-age) as this is the simplest formula to understand. As noted by the Reviewer, this approach is aligned with our Fair principle. Based on our experience, we decided that it would not be feasible to determine and utilise



individualised heart rate targets in the large number of schools included in our study. Finally, our B2L smartphone app uses age-predicted heart rate.

While collecting heart data during the shuttle run test was considered, due to logistical barriers (i.e., time constraints) we decided that it would not be a feasible to use heart rate monitors with a large group of participants. Moreover, our approach has ecological validity and is consistent with the normal practice for PACER assessment in the field.

2.13: Discussion. Line 468 – Given the length of the manuscript you might like to reconsider re-stating the Beets theory here as it is already discussed in the theoretical framework at line 189.

As suggested by the Reviewer, the two sentences describing Beets and colleagues' theory have been removed from the discussion.

2.14: References – Some references are missing journal volume and page number details, e.g. ref #s 9, 72 for instance.

We thank the Reviewer for noticing this. These references have now been amended.

2.15: Tables – Please add table legends defining abbreviations, such as B2L and SAAFE, so the tables can stand-alone.

We have now included legends for all tables.

Reviewer: 3

3.1: Abstract: First use of abbreviation (B2L) not defined. HIIT, which is a well-known abbreviation, is defined in the first sentence.

This has been addressed. We now state:

“Methods and analysis: The Burn 2 Learn (B2L) intervention will be evaluated using a two-arm parallel group cluster randomised controlled trial with allocation occurring at the school level (to treatment or wait-list control).”

3.2: Page 6: Cortisol may be elevated when subjects are in a low energy balance state (see Torstveit et al. 2018; Fahrenholtz et al. 2018; and Mountjoy et al. [RED-S] 2018), yet this protocol includes no strategy for assessing the state of energy balance (i.e., are they in a state of Relative Energy Deficiency in Sport [RED-S]?) at the time of the HIIT session. Including such an assessment strategy would help to better understand the potential benefits of HIIT in subjects who are in or not in a good energy balanced state. Performing HIIT when in a low energy balance state is likely to negatively influence the potential benefits of the exercise. Since subjects are from different schools, the HIIT may be performed at different times that influence energy balance. For instance, subjects performing HIIT shortly after lunch are likely to be in a better energy balance and achieve better outcomes, than those who perform HIIT just before lunch.

We thank the Reviewer for their valid point. Of note, our study is a large-scale implementation trial, which will involve more than 700 students and we will train teachers to deliver the HIIT sessions during their lessons. As such, we acknowledge that there will be variability in the timing of the sessions and students' acute responses to exercise. However, it is important to note that our study is focused on the chronic effects of exercise on stress, rather than the acute effects. In our study, hair cortisol concentrations will be used to represent participants' cumulative stress levels over the past two to three months. Consistent with cross stressor hypothesis, we hypothesize that increasing students' cardiorespiratory fitness will reduce their stress reactivity and contribute to lower self-

reported and objectively (i.e., hair cortisol concentrations) measured stress. Interestingly, a recent systematic review (see Mücke et al 2018, Sports Med)<sup>6</sup> identified a lack of experimental studies that had tested the acute effects of exercise on stress reactivity. In summary, additional studies are needed to examine the acute effects of exercise on stress in young people, yet this is beyond the scope of the current study.

3.3: Page 8: Delete the word "overall" as it is ill-defined and diminishes the power of the sentence.

As suggested by the Reviewer, the word "overall" has been removed from the sentence.

3.4: Page 12: This appears to be an appropriate protocol for randomization of Cohort 1 and Cohort 2 schools.

We thank the Reviewer for their comment.

3.5: Page 15: The body composition protocol fails to adequately assess the relative change in body fat/lean mass as a result of the protocol. BMI has been shown to be a poor measure for assessing relative body fatness in children. Suggest adding a validated strategy for assessing body composition.

This is a valid point raised by the Reviewer and we note this as a study limitation. However, we will continue to use BMI as a proxy for body composition for the following reasons. First, body composition is not a primary outcome in our study and alternate measures are generally more invasive (e.g., skinfolds) and/or expensive (e.g., DEXA). Second, based on pilot studies and the volume of HIIT prescribed in our current trial, we are unlikely to detect an improvement in body composition. Finally, our study has commenced and it is not possible to change the body composition measure.

3.6: Page 18: There is a nutritional component to psychological well-being that is not addressed. The brain is highly reliant on blood glucose for energy, and blood glucose can drop rapidly during the adolescent growth spurt and can also drop rapidly with physical activity if not actively supported with an appropriate nutritional protocol (eg. sipping on a sports beverage occasionally during activity, etc.).

We thank the Reviewer for their comment and acknowledge their concern. Similar to comment 3.2, we are only examining the chronic effect of exercise on psychological well-being. Although our intervention does not include a nutritional component, study participants will be instructed to stay hydrated during and after the HIIT sessions. Our research team have conducted a large number of health behaviour change interventions targeting adolescents, some of which have focused on dietary behaviour change. We have found that it is very difficult to change adolescents' dietary behaviours without altering the school food environment (e.g., school canteens), which is beyond the scope of our current trial.

3.7: Page 20: There is a formatting issue that makes the sentence unclear.

Thank you for the Reviewer for picking up on this error. It has now been addressed.

3.8: General: There is a strong relationship between nutritional state and the potential benefits that can be achieved from physical activity. Failure to account for this relationship may diminish the usefulness of the results.

Once again, we thank the Reviewer for their comment. From our experience, it is difficult to accurately assess dietary intake without using interviewer administered 24 hour food recalls. While this additional information may enhance the usefulness of our findings, our study already includes a large number of measures. Nutrition is not the focus of this intervention, and the large sample size and randomisation should deal with these differences to ensure it does not impact on our findings. In summary, we

believe that the assessment of nutritional status would place an unacceptable burden on schools and students.

#### References

1. Mahar MT, Welk GJ, Rowe DA. Estimation of aerobic fitness from PACER performance with and without body mass index. *Meas Phys Educ Exerc Sci* 2018;22(3):239-49. doi: 10.1080/1091367X.2018.1427590
2. Alberto PA, Troutman AC. *Applied Behavior Analysis for Teachers*. Australia: Pearson 2003.
3. Riley N, Lubans DR, Holmes K, et al. 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* 2016;13:198-206. doi: <http://dx.doi.org/10.1123/jpah.2015-0046>
4. Hardy LL, Mhrshahi S, Drayton BA, et al. NSW Schools Physical Activity and Nutrition Survey (SPANS) 2015: Full Report. In: NSW Department of Health, ed. Sydney, 2016.
5. White IR, Carpenter J, Horton NJ. Including all individuals is not enough: lessons for intention-to-treat analysis. *Clinical trials (London, England)* 2012;9(4):396-407. doi: 10.1177/1740774512450098 [published Online First: 2012/07/04]
6. Mucke M, Ludyga S, Colledge F, et al. Influence of regular physical activity and fitness on stress reactivity as measured with the trier social stress test protocol: A systematic review. *Sports medicine (Auckland, NZ)* 2018;48(11):2607-22. doi: 10.1007/s40279-018-0979-0 [published Online First: 2018/08/31]

#### VERSION 2 – REVIEW

<b>REVIEWER</b>	Stephen Malden University of Strathclyde, UK
<b>REVIEW RETURNED</b>	05-Feb-2019

<b>GENERAL COMMENTS</b>	The authors have addressed my comments and the paper is suitable for publication.
-------------------------	---

<b>REVIEWER</b>	Kate Ridley Flinders University Adelaide, Australia
<b>REVIEW RETURNED</b>	06-Mar-2019

<b>GENERAL COMMENTS</b>	The authors have provided considered and appropriate responses to my queries. I recommend that the manuscript is now ready for publication.
-------------------------	---