

## 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>	Physical activity and risk of behavioural and mental health disorders in Kindergarten children: analysis of a series of cross-sectional complete enumeration (census) surveys
<b>AUTHORS</b>	O'Brien, Kathleen; Agostino, Jason; Ciszek, Karen; Douglas, Kirsty

### VERSION 1 – REVIEW

<b>REVIEWER</b>	Francesco Sera London School of Hygiene and tropical Medicine, UK
<b>REVIEW RETURNED</b>	04-Nov-2019

<b>GENERAL COMMENTS</b>	<p>In this paper the authors analyse the association between physical activity and behavioural and mental health disorders (as measured by SDQ) in ~6 years old children in Australia. To evaluate the association the authors used information gathered from a series of cross-sectional surveys.</p> <p>The paper has same: the cross-sectional design doesn't allow to evaluate the direction of the association between PA and SDQ, and PA is measured through a questionnaire and not using objective measurements. The paper has also same strength: the sample size is adequate to evaluate the association and the information comes from a population based survey, moreover the paper is well written with a clear description of the results and discussion section.</p> <p>I have several major points related to method used in this research paper:</p> <ol style="list-style-type: none"><li>1. The description of the questionnaire used to measure PA should be more detailed, e.g. the authors could report the question used as supplementary material. More detail is also need on how the collected information were combined to determine the average daily PA level.</li><li>2. More detail is needed also for the KHC survey. Are kindergartens sampled as primary sample units? Are children sampled within a kindergarten? In their analysis the authors didn't consider the sampling design, and the possible clustering of children within a kindergarten.</li><li>3. The three sweeps of the KHC survey were simply pooled removing possible duplicates. I would consider also an indicator in the main model to model the sweep as a fixed effect.</li></ol> <p>I have the following minor points:</p> <ol style="list-style-type: none"><li>1. In their main model, which results are reported in table 3, the authors consider PA in a continuous scale. In order to assess also possible non-linearity I would consider a</li></ol>
-------------------------	--

	<p>sensitivity analysis on which PA is modelled as categorical variable (e.g. using categories reported in figure 2: &lt;30 min, 30-60 mins, 60-90 mins, &gt;90 mins).</p> <p>2. In table 3 I would report the Odds Ratios also when they are not significant. These ORs can be used in subsequent meta-analysis.</p> <p>3. In the introduction and in the discussion the authors didn't consider recent findings from the UK Millenium Cohort Study:</p> <p>a. Ahn, Jane V; Sera, Francesco; Cummins, Steven; Flouri, Eirini; Associations between objectively measured physical activity and later mental health outcomes in children: findings from the UK Millennium Cohort Study. J Epidemiol Community Health, 2018. b. Griffiths, Lucy; Geraci, Marco; Cortina-Borja, Mario; Sera, Francesco; Law, Catherine; Joshi, Heather; Ness, Andrew; Dezateux, Carol; Associations between children's behavioural and emotional development and objectively measured physical activity and sedentary time: Findings from the UK Millennium Cohort Study. Longitudinal and Life Course Studies, 2016.</p>
--	--

<b>REVIEWER</b>	A/Prof Hayley Christian University of Western Australia Australia
<b>REVIEW RETURNED</b>	05-Nov-2019

<b>GENERAL COMMENTS</b>	<p>Abstract – throughout be clear about what the SDQ measures ie risk of MH disorders or risk of MH and behavioural disorders? Objectives – PA more appropriate term for early childhood than exercise Design and title – enumeration surveys needs defining/explaining Outcome measures – how was the data collected – parent report, teacher report? Were demographic characteristics considered an outcome measure? Results – not necessary to include the completion numbers again here. High level meeting PA recommendations – what were these? The results refer to predictors – this is not a longitudinal study – it states in the limitations that it is a repeated cross-sectional. Known socio-demographic correlates more appropriate to refer to here. Conclusions – it is not clear what this study adds to the literature.</p> <p>Introduction PA vs exercise terminology and consistency needs checking throughout the paper. PA more appropriate for this young age group. The PA recs should try and be more broader than the Australian state in which the study was conducted ie Australian and other country specific (and common) PA recommendations and initiatives. Since the focus of the study is on early childhood the Introduction should also focus on this age group. Pls see systematic reviews on the topic that informed the 2017 Australian 24 hour movement guidelines for the early years (special issue BMC Public Health). Page 4, line 31, pls update as per current Australian 24 hour movement guidelines. I feel the authors have missed a significant amount of literature on PA and mental, social, behavioural and developmental outcomes in young children.</p>
-------------------------	---

	<p>Methods</p> <ul style="list-style-type: none"> <li>- Pls provide more details on the measures used to collect parent report PA levels of children – what measures were used and were they validated and reliable?</li> </ul> <p>The protocol for measuring height and weight could be included. How were participants recruited? How was the data collected? Was it compulsory? Online?</p> <p>What adjustment variables were included in the analyses?</p> <p>Results</p> <p>This is a repeated cross-sectional study so variables cannot be predictors – only correlates.</p> <p>These findings are similar to other national based data sets and findings published from the Longitudinal Study of Australian Children and Australian Early Development Census. - Table 3 – what confounders were adjusted for in each of the models? Eg PA should adjust for child sex, age, SES and Indigenous.</p> <p>Discussion</p> <p>Pls discuss your findings in relation to the Australian National Health Survey results and in particular in relation to objective vs parent report measures of PA.</p> <p>I am not sure that the analyses did examine if higher PA was associated with greater SDQ scores stratified by sex – this would be required to confirm the statements on page 9, line 9. -</p> <p>This is a very descriptive study and the first few paragraphs of the Discussion do not compare and contrast the findings with what we already know about the correlates of socialemotional problems in children as measured by the SDQ. -</p> <p>Discussion refers to some highly influential reports and publications but doesn't seem to be able to relate them specifically to the study's descriptive findings.</p> <p>Limitations – parent report PA, individual level measures of socio-demographic variables</p> <p>Conclusions</p> <p>Could the lack of findings with PA and the SDQ be due to the measure and cut offs used focussed on 'clinically significant' vs. at risk at a population level?</p> <p>It is not clear where the results support the statement that increased PA was associated with less peer problems and better prosocial behaviour?</p>
--	--

## VERSION 1 – AUTHOR RESPONSE

### RESPONSE TO REVIEWER COMMENTS

The authors thank the Reviewers for taking the time to read our paper and provide detailed and constructive feedback.

#### REVIEWER #1

In this paper the authors analyse the association between physical activity and behavioural and mental health disorders (as measured by SDQ) in ~6 years old children in Australia. To evaluate the association the authors used information gathered from a series of cross-sectional surveys. The paper has some limitations: the cross-sectional design doesn't allow to evaluate the direction of the association between PA and SDQ, and PA is measured through a questionnaire and not using objective measurements. The paper has also some strengths: the sample size is adequate to evaluate the association and the information comes from a population based survey, moreover the paper is well written with a clear description of the results and discussion section.

The authors thank the Reviewer for taking the time to read our paper and provide feedback. I have several major points related to method used in this research paper:

1. The description of the questionnaire used to measure PA should be more detailed, e.g. the authors could report the question used as supplementary material. More detail is also needed on how the collected information were combined to determine the average daily PA level. We have included information about the survey and the survey instrument, including the questions used to measure physical activity as supplementary material, as recommended. (Supplementary Files 1 & 2)

2. More detail is needed also for the KHC survey. Are kindergartens sampled as primary sample units? Are children sampled within a kindergarten? In their analysis the authors didn't consider the sampling design, and the possible clustering of children within a kindergarten. More information and clarity about the KHC has been included. The KHC is a complete enumeration survey, which means that all children enrolled in ACT Kindergartens are invited to participate in the survey. Participation is voluntary and is open to all eligible children. As the survey is a census there is no sampling error associated with it.

The participants section of the abstract has been modified as follows:

Participants: "All children enrolled in their first year of full-time primary education

(Kindergarten) were invited to participate. Of the 16,662 eligible Kindergarten children, 15,040 completed the survey."

The following text is included in the Methods:

"The KHC is an annual cross-sectional survey of all children in the ACT in their first year of full-time primary education (Kindergarten). All children enrolled in ACT Kindergartens are invited and are eligible to participate in the survey; participation is voluntary."

We reviewed response rates from individual schools, and of the 105 schools, all but one had a response rate over 75%, and all a response rate over 65%. The median response rate was 93%. As

response rates are high across the population, we are confident that clustering is not of concern. The authors did consider the sampling design; however, we acknowledge that more explicit description of this makes this clearer. An additional line has been added to the description of statistical analysis.

“The KHC is a census and thus there is no sampling error associated with it.”

3. The three sweeps of the KHC survey were simply pooled removing possible duplicates. I would consider also an indicator in the main model to model the sweep as a fixed effect.

We were not certain what the reviewer meant by this statement. We are confident there are no duplicate subjects across years, as these were searched for and deleted. If the reviewer was suggesting we look for a cohort effect that was different for each year, we have addressed this below by re-running the models.

The models were re-run as suggested using an indicator for the school year / cohort. Interestingly this indicator was significant in the models, but had a negligible effect on the model estimates of odds ratios. The examples are show in the table below (columns 2-3). Iterations of model for total difficulties

Independent variables	Original model	With year flag	Categorical PA	Categorical PA with year flag
Sex	1.80 (1.49-2.17)	1.79 (1.49-2.16)	1.80 (1.49-2.17)	1.80 (1.49-2.17)
Indigenous status	2.72 (1.78-4.16)	2.68 (1.75-4.10)	2.69 (1.76-4.11)	2.65 (1.73-4.05)
SES Q1	1.86 (1.38-2.50)	1.87 (1.40-2.51)	1.85 (1.37-2.48)	1.86 (1.38-2.50)
SES Q2	1.50 (1.11-2.03)	1.50 (1.11-2.02)	1.50 (1.11-2.02)	1.49 (1.10-2.01)
SES Q3	1.30 (0.96-1.77)	1.30 (0.96-1.76)	1.29 (0.95-1.75)	1.29 (0.95-1.75)
SES Q4	1.29 (0.95-1.76)	1.29 (0.95-1.76)	1.28 (0.94-1.75)	1.29 (0.95-1.75)
Physical activity (continuous)	0.96 (0.88-1.04)	0.96 (0.88-1.05)		
PA cat 2 (30-60)			0.95 (0.71-1.27)	0.95 (0.71-1.27)
PA cat 3 (60-90)			0.67 (0.49-0.91)	0.66 (0.49-0.90)
PA cat 4 (90+)			0.93 (0.71-1.22)	0.93 (0.71-1.22)
Underweight	1.05 (0.72-1.53)	1.06 (0.73-1.54)	1.06 (0.73-1.54)	1.06 (0.73-1.55)

Overweight	1.05 (0.79-1.38)	1.05 (0.80-1.39)	1.04 (0.79-1.37)	1.05 (0.80-1.38)
Obese	1.86 (1.27-2.72)	1.88 (1.28-2.73)	1.85 (1.27-2.70)	1.86 (1.27-2.72)
Age	1.03 (1.00-1.05)	1.03 (1.00-1.05)	1.03 (1.00-1.05)	1.03 (1.00-1.05)
Year Cat 2		0.75 (0.61-0.94)		0.75 (0.61-0.94)
Year Cat 3		0.77 (0.62-0.95)		0.76 (0.61-0.93)

As the inclusion of the cohort indicator had a negligible effect on the odds ratios, we have not included these results in the main paper, however we are happy to include in supplementary tables if requested by the editorial team.

I have the following minor points:

1. In their main model, which results are reported in table 3, the authors consider PA in a continuous scale. In order to assess also possible non-linearity I would consider a sensitivity analysis on which PA is modelled as categorical variable (e.g. using categories reported in figure 2: <30 min, 30-60 mins, 60-90 mins, >90 mins).

Analysis with physical activity categorised, rather than continuous, was undertaken. The table above shows the impact of this for the Total difficulties model (columns 4-5). This had a minimal effect on the model estimates of odds ratios for the other independent variables. Interestingly, physical activity participation of 60-90 minutes per day had reduced odds of being high risk for total difficulties (as compared to <30 minutes per day), whereas the other categories were not significant.

The authors have incorporated the sensitivity analysis into the manuscript (Supplementary File 3 - Sensitivity analysis using categorical physical activity). And additional text has been added to the methods:

“A sensitivity analysis was conducted with categories of physical activity participation (in lieu of continuous data).”

The results are referred to in the body of the manuscript, with tables included as supplementary material.

“Logistic regression analyses for Total difficulties and the SDQ subscales where physical activity was included as a categorical variable were done as part of a sensitivity analysis (Supplementary File 3). Odds ratios for independent variables, other than physical activity, showed only minor differences from those in the models where physical activity was included as a continuous variable. For Total difficulties and hyperactivity & inattention, 60 to <90 minutes of daily physical activity was significant in the models (baseline, <30 minutes daily). For prosocial behaviour, all categories were significant compared to baseline. Physical activity was not significant for the remaining SDQ subscales.” 2. In table 3 I would report the Odds Ratios also when they are not significant. These ORs can be used in subsequent meta-analysis.

Table 3 has been updated as advised to include all odds ratios, including those that are not statistically significant.

3. In the introduction and in the discussion the authors didn't consider recent findings from the UK Millenium Cohort Study:

a. Ahn, Jane V; Sera, Francesco; Cummins, Steven; Flouri, Eirini; Associations between objectively measured physical activity and later mental health outcomes in children: findings from the UK Millennium Cohort Study. *J Epidemiol Community Health*, 2018.

b. Griffiths, Lucy; Geraci, Marco; Cortina-Borja, Mario; Sera, Francesco; Law, Catherine; Joshi, Heather; Ness, Andrew; Dezauteux, Carol; Associations between children's behavioural and emotional development and objectively measured physical activity and sedentary time: Findings from the UK Millennium Cohort Study. *Longitudinal and Life Course Studies*, 2016.

The authors thank the reviewer for alerting them to these studies. Griffiths et al. 2016 has been referred to in the introduction and discussion.

## REVIEWER #2

The authors thank the Reviewer for taking the time to read our paper and providing detailed feedback. Abstract – throughout be clear about what the SDQ measures ie risk of MH disorders or risk of MH and behavioural disorders?

The document has been proof read to ensure consistency.

Objectives – PA more appropriate term for early childhood than exercise

The manuscript has been updated to use the term physical activity exclusively, rather than alternating with the term exercise.

Design and title – enumeration surveys needs defining/explaining

The phrase “complete enumeration” has had “(census)” added to it for clarity.

Title: “Physical activity and risk of behavioural and mental health disorders in Kindergarten children: analysis of a series of cross-sectional complete enumeration (census) surveys” Design: “A series of cross-sectional complete enumeration (census) surveys.” The sections on participants has also been revised to clarify this:

Participants: “All children enrolled in their first year of full-time primary education

(Kindergarten) were invited to participate. Of the 16,662 eligible Kindergarten children, 15,040 completed the survey.”

Outcome measures – how was the data collected – parent report, teacher report?

Were demographic characteristics considered an outcome measure?

This text was revised for clarity and to address reviewer comments.

Outcome measures: “Average daily physical activity participation derived from parent-reported data.

Prevalence of risk of behavioural and mental health disorders derived from the parent-completed Strengths and Difficulties Questionnaire (SDQ). Characteristics associated with SDQ Total difficulties and subscales (emotional difficulties, conduct problems, hyperactivity & inattention, peer problems, and prosocial behaviour) were estimated using logistic regression.”

However, due to the word limit on the abstract, we needed to exclude some key information.

Outcome measures: “Average daily physical activity participation and prevalence of risk of behavioural and mental health disorders derived from parent-reported data and the Strengths and

Difficulties Questionnaire (SDQ). Characteristics associated with SDQ Total difficulties and subscales were estimated using logistic regression. “ Results – not necessary to include the completion numbers again here.

These have been deleted.

High level meeting PA recommendations – what were these?

Text added in to clarify:

“8,340 (61.7%) children met physical activity targets (60 minutes or more daily) ...”

“Average daily physical activity was not significant, despite the highest levels of physical activity (90 minutes or more daily) being reported ... “

The results refer to predictors – this is not a longitudinal study – it states in the limitations that it is a repeated cross-sectional.

The abstract terminology has been revised as follows:

“... variables associated with high risk of behavioural and mental health disorders ... “ Known sociodemographic correlates more appropriate to refer to here. The abstract terminology has been revised as follows:

“Known sociodemographic correlates were also those variables associated with high risk ... “  
Conclusions – it is not clear what this study adds to the literature.

The conclusion has been edited to better highlight what the study adds to the literature.

“Our study provides comprehensive cross-sectional data on the relationship between physical activity participation and the risk of behavioural and mental health disorders in a large cohort of early school aged Australian children. Aboriginal and Torres Strait Islander children, boys, and those from the most disadvantaged socioeconomic group were at greatest risk of clinically significant behavioural and mental health disorders.”

Introduction

- PA vs exercise terminology and consistency needs checking throughout the paper. PA more appropriate for this young age group.

Terminology has been revised to refer physical activity throughout, rather than exercise.

- The PA recs should try and be more broader than the Australian state in which the study was conducted ie Australian and other country specific (and common) PA recommendations and initiatives.

The text has been modified as follows:

Physical inactivity is associated with many potentially preventable chronic diseases[1] and childhood is a key time for increasing participation in physical activity.[2] The benefits of physical activity start in childhood with protective effects on cardiovascular and metabolic health.[3,4] The importance of physical activity for school-aged children has been recognised by the World Health Organization, and within Australia through the Australian 24-hour movement guidelines for children and young people (5 to 17 years).[5,6]

- Since the focus of the study is on early childhood the Introduction should also focus on this age group.

The introduction has been modified to focus on children around five years of age.



“The evidence on the association between physical activity and mental wellbeing and behavioural disorders in young children is mixed. Several studies have associated physical activity participation with lower rates of depression and anxiety, and improved self-esteem.[7-8] However, mixed and contradictory results have been found and there is a paucity of information for younger children.[7,9,10] In their systematic review of early childhood physical activity and psychosocial wellbeing, Hinkley et al.[11] found supportive, null and even adverse associations between physical activity and a range of psychosocial wellbeing and behavioural outcomes. More recent systematic reviews continue to find mixed evidence. Carson et al. [12] showed a favourable relationship between physical activity and psychosocial health in preschool aged children based on an experimental study, while observational studies gave mixed results. Poitras et al. [13] assessed the relationship between objectively measured physical activity and indicators such as behavioural conduct/prosocial behaviour, psychological distress and self-esteem in 5-17-year-olds, however none of the included studies were for primary school-aged children. Nevertheless, physical activity may have a role in mitigating the increasing burden of mental health disorders in children. Anxiety, autism spectrum disorder, conduct disorder, and depressive disorders ranked among the six largest contributors to the burden of disease for Australian boys and girls aged 5-14 years in 2011.[1] The 2013-14 Australian Child and Adolescent Survey of Mental Health and Wellbeing (Young Minds Matter) survey reported a mental disorder in 14% of boys and 7% of girls in Kindergarten or pre-primary.[14] “

- Pls see systematic reviews on the topic that informed the 2017 Australian 24 hour movement guidelines for the early years (special issue BMC Public Health). The authors appreciate being directed to these resources. We have included references to Carson et al. 2017, Poitras et al. 2016, and 24-Hour Movement Guidelines Leadership Group 2016. - Page 4, line 31, pls update as per current Australian 24 hour movement guidelines. The manuscript has been revised to refer to 24-hour movement guidelines rather than physical activity guidelines throughout.

The authors appreciate being directed to the revised Australian 24-hour movement guidelines. We note that there are two potentially relevant guidelines: early years (birth to 5 years); children and young people (5 to 17 years). We consider the latter of more relevance to our study, as this is aimed at children in primary and secondary school, and the children in our study are all enrolled in their first year of full-time primary school. We note that the Australian 24-hour movement guidelines have used Canadian research in their development, and that research for children in primary school has also been updated by the Canadian Society for Exercise Physiology. These findings have been included in the manuscript.

- I feel the authors have missed a significant amount of literature on PA and mental, social, behavioural and developmental outcomes in young children. The authors acknowledge that significant research for pre-school aged children has been completed in recent years. Our study, however, is on children in their first year of full time primary education; that is, the school year following pre-school. We have added references to the body of work that relates to our study population as noted above.

## Methods

- Pls provide more details on the measures used to collect parent report PA levels of children – what measures were used and were they validated and reliable? The KHC survey instrument has been added as supplementary material (Supplementary File 2). The questions used to collect information on physical activity are included within the survey instrument. - The protocol for measuring height and weight could be included.- How were participants recruited? How was the data collected? Was it compulsory? Online?

This study was based on an existing data collection. Study protocols have therefore been included as supplementary material.

## Supplement 1: Kindergarten Health Check methods

The Kindergarten Health Check (KHC) is conducted annually in the Australian Capital Territory (ACT). See: <https://health.act.gov.au/services-and-programs/women-youth-andchildren/childrenand youth/school-health>

All children in their first year of full-time primary education (Kindergarten) in the ACT are invited to participate. Participation is voluntary. Paper survey forms are distributed to parents/guardians of eligible children via their schools in early February of the school year. Parent-completed questionnaires are collected by schools a few weeks later.

Physical health checks are conducted by School Health Nurses between April and October of that year for all children whose parents/guardians' consent to their participation. The physical health check includes body measurements (height, weight), vision check and hearing check. Results are recorded on a paper survey form. The KHC is conducted using best practice protocols as informed by recent literature, including protocols for measuring height and weight.

Denominator data is derived from The ACT Education Directorate census data.

The same survey instrument was used for 2014-2016. A copy of the 2016 survey instrument is included as Supplementary File 2.

- What adjustment variables were included in the analyses?  
Independent [adjustment] variables are described under "Statistical analysis". Page 4, lines 41-48 of original manuscript Results

- This is a repeated cross-sectional study so variables cannot be predictors – only correlates. This is understood and discussed the limitations, however we have endeavoured to ensure the language throughout the manuscript is clearer on this issue.
- These findings are similar to other national based data sets and findings published from the Longitudinal Study of Australian Children and Australian Early Development Census.

We have reviewed LSAC and our results apply to a different age group, specifically looking at PA and mental health and behavioural difficulties; the longitudinal study assesses related issues, such as screen time and emotional wellbeing, and PA in older children than our cohort.

Results from the AEDC have been included in the discussion.

"The Australian Early Development Census (AEDC) is a three-yearly census of children in their first year of full-time primary education. Data from the 2018 AEDC found similar patterns to our study. Boys were more likely to be developmentally vulnerable than girls; including the social competence and emotional maturity domains.[22] Aboriginal and Torres Strait Islander children were more likely to be developmentally vulnerable than other children, as were children from the most disadvantaged socioeconomic groups."

- Table 3 – what confounders were adjusted for in each of the models? Eg PA should adjust for child sex, age, SES and Indigenous.

Table 3 shows the variables included in each of the models: sex, Indigenous status, quintile of relative socioeconomic advantage [IRSD], physical activity, BMI category, age. This is also described in the methods under "statistical analysis".

## Discussion

- Pls discuss your findings in relation to the Australian National Health Survey results and in particular in relation to objective vs parent report measures of PA.

A paragraph including these data has been added to the discussion. However, my reading of the explanatory notes is that the data for children aged 5-8 years (without high RSEs) is also self-reported.

The pedometer data has high RSEs.

“The 2011-12 AHS found similar proportions of boys (36%) and girls (35%) aged 5-8 years met the recommended 60 minutes or more of physical activity on all seven days prior to interview.[15] However similar to our study, boys undertook more daily exercise. Boys aged 5-17 years spent an average of 78 minutes in moderate to vigorous activity daily, compared to 67 minutes for girls. There were no statistically significant differences in the proportion of children aged 5-17 years meeting physical activity targets based on socioeconomic status (IRSD).”

A comment about measured versus parent-reported PA data has been included in the discussion. “While not subject to sampling error, the data are subject to non-sampling error. For example, the physical activity measure is derived from parent-reported data. As with anthropometric data, measured, objective, physical activity data is more accurate than parent-reported activity.[30]” - I am not sure that the analyses did examine if higher PA was associated with greater SDQ scores stratified by sex – this would be required to confirm the statements on page 9, line 9. This statement is based on the cross-sectional results, where physical activity participation was analysed by sex (Figure 2), and SDQ scores were separately analysed by sex (Table 2).

- This is a very descriptive study and the first few paragraphs of the Discussion do not compare and contrast the findings with what we already know about the correlates of social-emotional problems in children as measured by the SDQ.
- Discussion refers to some highly influential reports and publications but doesn't seem to be able to relate them specifically to the study's descriptive findings.

We have endeavoured to edit the discussion to address these concerns. For example:

“Our results are consistent with SDQ data for Australian children aged 4-12 years; higher rates of high risk Total difficulties were seen for boys compared to girls (12.7% versus 7.7%) and children in the most disadvantaged socioeconomic groups (15.6% in the lowest SES quintile versus 7.2% in the highest SES quintile).[20] Aboriginal and Torres Strait Islander children have also been shown elsewhere to have higher SDQ scores than the Australian average.[21]” - Limitations

– parent report PA,

We agree that this [parent report PA] is an important point, and have added explicit text in the discussion as noted above.

individual level measures of socio-demographic variables

The discussion includes paragraphs addressing sociodemographic disadvantage being area-based rather than individual. Please see page 10, line 55 to page 11, line 12 in the original manuscript.

Conclusions

- Could the lack of findings with PA and the SDQ be due to the measure and cut offs used focussed on ‘clinically significant’ vs. at risk at a population level?  
The authors agree this could be an explanatory factor, and have included some text to that effect in the discussion.

“We undertook the logistic regression using standard categories for degree of risk from the SDQ. It is possible that different relationships may have been observed if the raw scores were used, as was done in the study by Griffiths et al.[9]”

- It is not clear where the results support the statement that increased PA was associated with less peer problems and better prosocial behaviour?

The odds ratio for physical activity participation was 0.78 (0.62-0.98) for being high risk for peer problems, and 0.78 (0.70-0.87) for being high risk for prosocial behaviour. We have rephrased this for clarity in the conclusion.

“Lower levels of physical activity were associated with being at high risk for peer relationship problems or prosocial behaviour difficulties.”

### VERSION 2 – REVIEW

<b>REVIEWER</b>	Francesco Sera London School of Hygiene & Tropical Medicine London UK
<b>REVIEW RETURNED</b>	02-Jan-2020

<b>GENERAL COMMENTS</b>	<p>I think that overall the authors have answered positively to the points I rose in my previous review.</p> <p>Overall the manuscript is well written. and has improved from the original submission, and I think it could give a contribution to the literature.</p> <p>I suggest removing the sentence: "The KHC is a census and thus there is no sampling error associated with it.". In fact, this sentence is contrast with the inferential procedures used by the authors. Moreover, even if the original study design was a census, the population analysed can be considered a sample under the missing completely a random hypothesis.</p> <p>As a final remark, I disagree with the authors when they state that clustering is not a concern. Clustering is not related with the response rate within a school, but with the tendency of the outcome to be correlated within schools, and has an effect on the standard error of the coefficients (especially on contextual variables as the SES indicator used by the authors).</p>
-------------------------	---

### VERSION 2 – AUTHOR RESPONSE

#### RESPONSE TO REVIEWER COMMENTS

The authors thank the Reviewer once again for taking the time to read our paper and provide constructive feedback.

#### REVIEWER #1

I think that overall the authors have answered positively to the points I rose in my previous review.

Overall the manuscript is well written. and has improved from the original submission, and I think it could give a contribution to the literature.

The authors thank the reviewer for their feedback which has helped us improve our submission. We are particularly appreciative of the reviewer taking the time to explain their interpretation of the methods.

I suggest removing the sentence: "The KHC is a census and thus there is no sampling error associated with it.". In fact, this sentence is contrast with the inferential procedures used by the authors. Moreover, even if the original study design was a census, the population analysed can be considered a sample under the missing completely at random hypothesis.

As a final remark, I disagree with the authors when they state that clustering is not a concern. Clustering is not related with the response rate within a school, but with the tendency of the outcome to be correlated within schools, and has an effect on the standard error of the coefficients (especially on contextual variables as the SES indicator used by the authors).

The authors thank the reviewer for their feedback and in particular for articulating their thoughts.

While we do consider the KHC subject to non-response error rather than sampling error per se, we agree that the study population could be thought of as a sample under the Missing Completely at Random approach. And acknowledge that if clustering occurs within schools that this may affect error estimates.

We have amended the text as follows:

The KHC data are subject to non-response and other non-sampling error. All Kindergarten children in all ACT schools were invited to participate, with an overall response rate of 90.6%. The characteristics of children who did not participate in the survey may have been different to those that did. As the response rates were high, any effects of non-response were considered unlikely to be significant. If outcomes were correlated within schools, this may cause clustering which would impact error estimates. For the purposes of our study, this was considered unlikely to be significant. All data are subject to non-sampling error. An example of this is the physical activity measure, which is derived from parent-reported data. As with anthropometric data, measured and objective, physical activity data is more accurate than parent-reported activity.[30]