

PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form ([see an example](#)) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below. Some articles will have been accepted based in part or entirely on reviews undertaken for other BMJ Group journals. These will be reproduced where possible.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Behavioural and weight status outcomes from an exploratory trial of the Healthy Lifestyles Programme (HeLP): A novel school-based obesity prevention programme
AUTHORS	Jennifer J Lloyd, Katrina M Wyatt and Siobhan Creanor

VERSION 1 - REVIEW

REVIEWER	Vera Araujo-Soares, PhD Senior Lecturer in Health Psychology Institute of Health & Society
REVIEW RETURNED	10/10/2011

THE STUDY	This is not a patient study hence the answer of NO to this question. More detail should be provided on the intervention. More detail should be provided on randomisation and on blinding of assessors.
REPORTING & ETHICS	Some of the CONSORT sections have not been provided like trial registration. The authors should describe ethics procedures in detail.
GENERAL COMMENTS	<p>ID bmjopen-2011-000390</p> <p>"Behavioural and weight status outcomes from an exploratory trial of the Healthy Lifestyles Programme (HeLP): A novel school-based obesity prevention programme."</p> <p>This is a very clear and well written manuscript that presents the results of a novel school-based obesity prevention programme on behavioural and weight outcomes in English primary school children. This manuscript describes the results of phase 3 of the MRC guidance on the development and evaluation of complex interventions.</p> <p>Please introduce abbreviations in the abstract (e.g., sds).</p> <p>In the methods section (page 8) more detail should be provided about the randomization method.</p> <p>In the intervention section the authors name the theory used to develop the intervention programme (Information, Motivation and Behavioural Skills Model-IMB). It would be worthwhile to better describe the model and its core assumptions. Also, reference to empirical evidence for this model should be provided.</p> <p>Figure 1 is extremely unclear. The fact that the term 'mediating variables' appears in three boxes is unhelpful. I suggest reconsidering this figure and revising it in a way that more clearly describes the relationships between key processes and measures in</p>

	<p>this study. You might find Hardeman et al (2005; Health Edu Res 20, 6, 676-687) helpful in this regard.</p> <p>On page 9 the authors discuss behaviour change techniques used within the intervention. It would be helpful to relate this to explicit published taxonomies of behaviour change techniques for example by Abraham & Michie (2008 – Health Psychology) or Michie et al (in press Psychology and Health). It would also be helpful to clarify in Table 1 the delivery of the intervention according to the criteria suggested by KW Davidson et al 2003; Annals of Behavioral Medicine 26; 3, 161-171). For example frequency and duration of newsletters, activity workshops, the PSHE lessons and the drama are unclear. In addition, it would be ideal to provide access to the intervention manuals as online supplement. These suggestions are important in making the intervention replicable, and to allow evidence synthesis based on a clear understanding of the details of content and delivery of this highly interesting intervention.</p> <p>On page 10, please explain/define the term “literacy groups”. Please when describing the accelerometer used inform readers if this equipment was water proof or water resistant.</p> <p>In the results section on page 13 the authors state that the mean sedentary time for all children was 16.2 hours. Is this including sleeping time or timing sitting during classes? Also please specify the number of participants in your sample for which you have valid accelerometer data and insert this on Table 2.</p> <p>General points:</p> <p>It is stated that the intervention presented in this manuscript has been developed to allow for flexibility and adaptation in accordance to specific needs. It would be good to provide more information on how this was done and how is this embedded in the intervention materials.</p> <p>No data is presented on fidelity. Have the authors considered this?</p> <p>Please clarify blinding procedures for outcome assessors.</p>
--	--

REVIEWER	<p>Dr Emma Lancashire Senior Research Fellow University of Birmingham UK</p> <p>I have no competing interests.</p>
REVIEW RETURNED	20/10/2011

THE STUDY	<p>Are the methods adequately described?</p> <p>Information should be provided in relation to how consent was obtained for the individuals taking part in the study</p> <p>Are the abstract/summary/key messages/limitations accurate?</p> <p>Abstract p2 lines 15 and 18 currently imply that follow up</p>
------------------	--

	<p>measurements were undertaken 18 and 24 months post intervention and not post baseline - reword to clarify this.</p> <p>p4 line 6 I think that the fact that the differences are relatively small should be acknowledged and would suggest inserting '(although the differences are relatively small)' after 'effects'.</p> <p>p4 line 11 I would suggest changing 'necessary' to 'justified'</p> <p>In the limitations I think the authors should acknowledge that although the schools were randomly allocated post baseline measurements to the control and intervention arms of the trial with only 4 schools and 202 children the numbers are too small to ensure an even balance of baseline characteristics between the 2 arms. (Although differences in the mean values for the outcomes presented are in general shown to be relatively small between the the 2 arms the control arm has 5.3% more obese children and 4.8% more children with bodyfat above or equal to the 85th centile than the intervention arm. The percentage of children eligible for a free school meal is also 4% higher in the control arm at baseline, as this is a proxy for socio economic status, it would suggest that the children in the control arm are from families with a slightly lower socioeconomic background and therefore likely to have less healthy eating habits.)</p> <p>Are the statistical methods appropriate: I am not able to comment sufficiently well on the appropriateness of the statistical methods and suggest the manuscript should be reviewed by a suitable statistician.</p> <p>It might be helpful to have the questionnaires used to measure food intake and TV viewing/screen time included as supplemental files.</p>
<p>RESULTS & CONCLUSIONS</p>	<p>Data are presented (Table 3) in terms of difference in the means of measured outcomes between intervention and control arms at follow up without considering what the difference between the two arms was at baseline. When the anthropometric data are presented the change in mean value from baseline to follow up in each arm is presented which makes sense. However for the behavioural data differences in the mean values between the 2 arms at follow up are presented with no reference to the difference at baseline. Without reporting on the difference in the mean values between the 2 arms at baseline it is not possible to determine whether differences observed at follow up between the intervention and control arm are in favour of the intervention. Although at follow up the mean values for the behaviours of the intervention group are healthier than those of the control group this could also have been true at baseline and not be as a result of the intervention. In fact when baseline differences are considered the changes observed in the differences between baseline and follow up do show changes indicating a shift towards healthier behaviours in the intervention arm when compared to the control arm. In terms of negative food markers there was no difference in the mean number of such markers between the intervention and control arm at baseline whereas at 18 months follow up children in the intervention schools had 0.69 less negative food markers. At baseline children at intervention schools consumed 0.1 more energy dense snacks whereas at follow up the children at intervention schools were consuming 0.28 less energy dense snacks than control schools. At baseline children from intervention schools consumed 0.3 less positive food markers than children from control schools whereas at follow up they were consuming 0.38 more. With</p>

	<p>regard to mean TV/screen time at baseline children from intervention schools spent 0.2 hours less than children from control schools in this type of activity this difference had increased to 0.41 hours less at follow up. Finally on average children from intervention schools spent 12.8 minutes less than children from control schools engaged in moderate to vigorous physical activity at baseline whereas at follow up children from intervention schools were spending on average 5.67 minutes more than children from control schools in such activities. I would have thought that presenting the data in this way would strengthen the argument. In addition to changing how the results are reported in the text, Table 2 should include the mean difference (Intervention minus Control) at baseline and Table 3 should include (in addition to the mean difference between the 2 arms at follow up) the values for the actual means for all outcomes at follow up for both the intervention and the control arms.</p>
GENERAL COMMENTS	<p>p6 line 3 'children's transport activity' I think this term needs further explanation</p> <p>p6 line 26 '4.5 hours in screen time a day' seems to me an odd phrase could this be reworded ? '4.5 hours a day looking at a screen'</p> <p>p8 line 6 change 'group' to 'class'</p> <p>p8 line 46 there is no space after the full stop</p> <p>p9 line 58/59 + p10 line 13/14 should it be 'the' adapted version or 'an' adapted version?</p> <p>p10 line 59 mentions that there was a 10% data entry check but no reference is made to the level of errors observed and in particular whether or not it was felt to be satisfactory</p> <p>p11 change 'each' to 'during the'</p> <p>p12 line 58/59 change 'years' to 'year olds'</p> <p>p12 line 60 (last line) says that percentage with no rules regarding amount of TV/screen time was higher in the control schools, this does not reflect Table 2 which reports 39.2% as the mean for intervention schools and 33.9% as the mean for control schools</p> <p>p13 line 42 before 'At 18 months' insert 'Compared to baseline, '</p> <p>p14 line 39 insert a comma after [35-37]</p> <p>p17 line 46 change 'warranted' to 'justified'</p>

REVIEWER	<p>George Ploubidis, London School of Hygiene and Tropical Medicine No competing interests</p>
REVIEW RETURNED	<p>18/01/2012</p>

GENERAL COMMENTS	<p>The manuscript is well written and the description of the exploratory cluster randomized trial is very clear. The authors do a very good job of summarising the existing literature and the statistical methods they</p>
-------------------------	---

	<p>employed are appropriate. However, I have some concerns, which I outline below:</p> <p>i) The authors mention that missing data were negligible. I believe it would be beneficial to the readers if they could provide more information on the proportion of missing data as this can be very influential on ITT analysis.</p> <p>ii) The authors present results on the school level and argue that the number of clusters is small (4) in order to adjust for individual level confounders. I agree that a random effects – multilevel- model is not feasible in this case, but a sensitivity analysis treating schools as a fixed effect and therefore controlling for individual level confounders would be beneficial to the readers. I am not suggesting that a fixed effects model should replace the cluster level analysis, but I think an individual level sensitivity analysis should be carried out.</p>
--	---

VERSION 1 – AUTHOR RESPONSE

Reviewer 1: Vera Araujo-Soares, PhD
 Senior Lecturer in Health Psychology
 Institute of Health & Society
 Faculty of Medical Sciences
 Newcastle University

1.1 Please introduce abbreviations in the abstract (e.g., sds).
 These have now been added in the abstract (p2).

1.2 In the methods section (page 8) more detail should be provided about the randomization method.
 More detail has now been added to the study design section in the methods (p8).

1.3 In the intervention section the authors name the theory used to develop the intervention programme (Information, Motivation and Behavioural Skills Model-IMB). It would be worthwhile to better describe the model and its core assumptions. Also, reference to empirical evidence for this model should be provided.
 This has now been added in the intervention section (p8).

1.4 Figure 1 is extremely unclear. The fact that the term ‘mediating variables appears in three boxes is unhelpful. I suggest reconsidering this figure and revising it in a way that more clearly describes the relationships between key processes and measures in this study. You might find Hardeman et al (2005; Health Edu Res 20, 6, 676-687) helpful in this regard.
 We agree that this figure is unclear and, as a result, have completely revised it and provided a clearer description of these key processes in the intervention section (p9).

1.5 On page 9 the authors discuss behaviour change techniques used within the intervention. It would be helpful to relate this to explicit published taxonomies of behaviour change techniques for example by Abraham & Michie (2008 – Health Psychology) or Michie et al (in press Psychology and Health). It would also be helpful to clarify in Table 1 the delivery of the intervention according to the criteria suggested by KW Davidson et al 2003; Annals of Behavioral Medicine 26; 3, 161-171). For example frequency and duration of newsletters, activity workshops, the PSHE lessons and the drama are unclear. In addition, it would be ideal to provide access to the intervention manuals as online supplement. These suggestions are important in making the intervention replicable, and to allow evidence synthesis based on a clear understanding of the details of content and delivery of this highly interesting intervention.

Reference to Abraham and Michie's taxonomies has been inserted as well as more information about the intervention (p 9). Table 1 now includes the duration as well as the frequency of the intervention components specified for each phase. We do not feel that providing access to intervention materials at this stage is appropriate as we are presenting the results of an exploratory trial not results from a definitive trial. We have very recently secured funding for the definitive trial of HeLP and we would consider providing access to complete manuals following completion of this large scale cluster RCT.

1.6 Please when describing the accelerometer used inform readers if this equipment was water proof or water resistant.

This information has now been inserted (p11).

1.7 In the results section on page 13 the authors state that the mean sedentary time for all children was 16.2 hours. Is this including sleeping time or timing sitting during classes? Also please specify the number of participants in your sample for which you have valid accelerometer data and insert this on Table 2.

This information has been inserted into the text (p12) when describing inclusion criteria for the analysis of accelerometry data rather than in table 2.

1.8 On page 10, please explain/define the term "literacy groups".

This information has been inserted (p11).

General points:

1.9 It is stated that the intervention presented in this manuscript has been developed to allow for flexibility and adaptation in accordance to specific needs. It would be good to provide more information on how this was done and how is this embedded in the intervention materials. Information on how this was done has been inserted in the intervention section (p9).

1.10 No data is presented on fidelity. Have the authors considered this?

Fidelity was considered in this exploratory trial and, as part of the process evaluation, (which will be reported fully in another paper) components of the intervention were observed and field notes taken. However, as this was an exploratory trial (to determine feasibility and acceptability of the intervention, trial processes and outcome measures) we do not feel that presenting this data is appropriate here. Fidelity to intervention using a comprehensive checking manual will be assessed in the definitive trial.

1.11 Please clarify blinding procedures for outcome assessors

We are not sure what additional information is required. We state that the assessor was 'independent, blind to allocation' meaning that they were not involved in the development or evaluation of HeLP and they did not know, when collecting the anthropometric data from the children, whether they were in an intervention or control school (p10).

1.12 Some of the CONSORT sections have not been provided like trial registration.

As this was an exploratory as opposed to a definitive trial, some of the usual CONSORT checks were not applicable, such as trial registration.

1.13 The authors should describe ethics procedures in detail.

We are happy to provide this information should the editors require it but this will add a further 150 words to the word count.

Reviewer 2: Dr Emma Lancashire

Senior Research Fellow
University of Birmingham
UK

2.1 Information should be provided in relation to how consent was obtained for the individuals taking part in the study

This information has been inserted in the study design (p8).

2.2 Abstract p2 lines 15 and 18 currently imply that follow up measurements were undertaken 18 and 24 months post intervention and not post baseline – reword to clarify this.

This has been amended as suggested (p2).

2.3 p4 line 6 I think that the fact that the differences are relatively small should be acknowledged and would suggest inserting '(although the differences are relatively small)' after 'effects'.

This has been inserted as suggested (p4).

2.4 p4 line 11 I would suggest changing 'necessary' to 'justified'

This has been inserted as suggested (p4).

2.5 In the limitations I think the authors should acknowledge that although the schools were randomly allocated post baseline measurements to the control and intervention arms of the trial with only 4 schools and 202 children the numbers are too small to ensure an even balance of baseline characteristics between the 2 arms. (Although differences in the mean values for the outcomes presented are in general shown to be relatively small between the the 2 arms the control arm has 5.3% more obese children and 4.8% more children with bodyfat above or equal to the 85th centile than the intervention arm. The percentage of children eligible for a free school meal is also 4% higher in the control arm at baseline, as this is a proxy for socio economic status, it would suggest that the children in the control arm are from families with a slightly lower socioeconomic background and therefore likely to have less healthy eating habits.)

We agree that if this study was the definitive trial to assess the effectiveness of HeLP, then not having a reasonable balance between intervention and control groups in key measures at baseline would be a limitation. However, as this is an exploratory trial, the main aim of which was to assess the feasibility and acceptability of the Programme and trial design to provide evidence of 'proof of concept' to justify funding the large scale cluster RCT, we do not feel that this imbalance is a limitation per se. What we do feel is important, though, is to acknowledge this imbalance (p13) as well as the fact that the estimates of the differences between intervention and control schools were imprecise (p17).

2.6 It might be helpful to have the questionnaires used to measure food intake and TV viewing/screen time included as supplemental files.

We can provide these as supplementary files if the editor thinks this is necessary for the presentation of results, however we have referenced them fully and the questionnaires are widely available.

2.7 Data are presented (Table 3) in terms of difference in the means of measured outcomes between intervention and control arms at follow up without considering what the difference between the two arms was at baseline. When the anthropometric data are presented the change in mean value from baseline to follow up in each arm is presented which makes sense.

We think the reviewer may have mis-understood the data presented in Table 3 as it is not the mean changes from baseline that are given, rather the mean difference between the two groups at each of the two follow up time points, for both the anthropometric and behavioural outcome measures. We do state in the statistical analysis section (p12) that only cluster level analyses were undertaken as there

were only 4 participating schools. Adjusting for individual-level covariates (e.g. baseline measures), therefore, would be statistically inappropriate.

However for the behavioural data differences in the mean values between the 2 arms at follow up are presented with no reference to the difference at baseline. Without reporting on the difference in the mean values between the 2 arms at baseline it is not possible to determine whether differences observed at follow up between the intervention and control arm are in favour of the intervention. Although at follow up the mean values for the behaviours of the intervention group are healthier than those of the control group this could also have been true at baseline and not be as a result of the intervention. In fact when baseline differences are considered the changes observed in the differences between baseline and follow up do show changes indicating a shift towards healthier behaviours in the intervention arm when compared to the control arm. In terms of negative food markers there was no difference in the mean number of such markers between the intervention and control arm at baseline whereas at 18 months follow up children in the intervention schools had 0.69 less negative food markers. At baseline children at intervention schools consumed 0.1 more energy dense snacks whereas at follow up the children at intervention schools were consuming 0.28 less energy dense snacks than control schools. At baseline children from intervention schools consumed 0.3 less positive food markers than children from control schools whereas at follow up they were consuming 0.38 more. With regard to mean TV/screen time at baseline children from intervention schools spent 0.2 hours less than children from control schools in this type of activity this difference had increased to 0.41 hours less at follow up. Finally on average children from intervention schools spent 12.8 minutes less than children from control schools engaged in moderate to vigorous physical activity at baseline whereas at follow up children from intervention schools were spending on average 5.67 minutes more than children from control schools in such activities.

I would have thought that presenting the data in this way would strengthen the argument. In addition to changing how the results are reported in the text, Table 2 should include the mean difference (Intervention minus Control) at baseline and Table 3 should include (in addition to the mean difference between the 2 arms at follow up) the values for the actual means for all outcomes at follow up for both the intervention and the control arms.

Table 2, which presents the baseline measures for the intervention and control groups separately (as well as overall), follows the generally accepted layout for presenting baseline data, as per the example in the CONSORT guidelines for cluster randomised trials. It is not good practice to formally test for differences between randomised groups at baseline (e.g. testing for baseline balance in clinical trials, Senn, *Statistics in Medicine*, 1994) and indeed the recommendation is to include important variables (e.g. baseline measures) in an analysis of covariance approach, regardless of their distribution (or statistical significance) at baseline. However, with such a small exploratory trial this was not possible (as detailed in the statistical analysis section (p12)).

We have intentionally not analysed any data as 'change from baseline' as it is widely accepted that from a statistical perspective, it is not best practice; instead it is more efficient to look for differences between groups at follow-up, using analysis of covariance approaches to adjust for baseline values. However, as stated above, this was not possible in this exploratory trial, but will be fully addressed in the analyses of the definitive study.

2.8 p6 line 3 'children's transport activity' I think this term needs further explanation
This has been inserted (p6).

2.9 p6 line 26 '4.5 hours in screen time a day' seems to me an odd phrase could this be reworded ?
'4.5 hours a day looking at a screen'
This has been amended as suggested (p6).

2.10 p8 line 6 change 'group' to 'class'
This has been amended as suggested (p8).

2.11 p8 line 46 there is no space after the full stop
This has been corrected (p8).

2.12 p9 line 58/59 + p10 line 13/14 should it be 'the' adapted version or 'an' adapted version?
These have been amended to 'an' adapted version as suggested (p10).

2.13 p10 line 59 mentions that there was a 10% data entry check but no reference is made to the level of errors observed and in particular whether or not it was felt to be satisfactory
Information of the outcome of data checking has been inserted (p12).

2.14 p11 change 'each' to 'during the'
This has been changed as suggested (p12).

2.15 p12 line 58/59 change 'years' to 'year olds'
This has been changed as suggested (p14).

2.16 p12 line 60 (last line) says that percentage with no rules regarding amount of TV/screen time was higher in the control schools, this does not reflect Table 2 which reports 39.2% as the mean for intervention schools and 33.9% as the mean for control schools
This was a mistake by us and has been amended accordingly (p14).

2.17 p13 line 42 before 'At 18 months' insert 'Compared to baseline,
This has not been amended as given there were only 4 schools only cluster level analyses were appropriate so we did not adjust for baseline values as stated in the analysis section (p12/13).

2.18 p14 line 39 insert a comma after [35-37]
This has been inserted (p16).

2.19 p17 line 46 change 'warranted' to 'justified'
This has been changed (p19).

Reviewer 3: George Ploubidis
London School of Hygiene and Tropical Medicine, Department of Epidemiology

3.1 The authors mention that missing data were negligible. I believe it would be beneficial to the readers if they could provide more information on the proportion of missing data as this can be very influential on ITT analysis.
Information on the numbers and proportions of missing data for the FIQ, TV viewing questionnaire and accelerometry (as well as for the anthropometric data) has been added to Figure 2.

3.2 The authors present results on the school level and argue that the number of clusters is small (4) in order to adjust for individual level confounders. I agree that a random effects – multilevel- model is not feasible in this case, but a sensitivity analysis treating schools as a fixed effect and therefore controlling for individual level confounders would be beneficial to the readers. I am not suggesting that a fixed effects model should replace the cluster level analysis, but I think an individual level sensitivity analysis should be carried out.

We thank the reviewer for this suggestion, which we have given due consideration. We feel, however, that as this is a small, exploratory study, with primary aims of testing feasibility, that to undertake and present more detailed analysis, even as a sensitivity analysis, could lead to over interpretation of the results found. However, if the editors feel that these additional analyses are required, we would be willing to undertake these (given sufficient time), although the inclusion of the results may somewhat lengthen the article.

VERSION 2 – REVIEW

REVIEWER	Vera Araujo-Soares, PhD Senior Lecturer in Health Psychology Institute of Health & Society Faculty of Medical Sciences No competing interests.
REVIEW RETURNED	21/02/2012

GENERAL COMMENTS	<p>Happy with the way the authors dealt with the comments made to a previous version of this manuscript.</p> <p>There are, nevertheless, some minor things that I would like to be considered by the authors.</p> <p>In response to point 1.2 the authors provide more details “to the study design section in the methods (p8).” On page 8 it is stated: “The class teacher... daily reminders...”. For how long and how many reminders were sent out? Oral reminders; for the children to transmit to their parents? Or written reminders?</p> <p>More important: in response to point 1.5. the authors state that they now refer to the taxonomy. The authors should refer to and more crucially use the new taxonomy: Michie, S., Ashford, S., Snihotta, F. F., Dombrowski, S. U., Bishop, A., French, D. P. (2011). A refined taxonomy of behaviour change techniques to help people change their physical activity and healthy eating behaviours: the CALO-RE taxonomy. <i>Psychol Health</i> 26(11), 1479-1498 doi:10.1080/08870446.2010.540664.</p> <p>It would also be key if the Behaviour Change Techniques listed on table 1 would map onto this taxonomy. It is not just a matter of referring to it, it is mostly a matter of using it for the sake of clarity within the description of the key BCT's within their intervention.</p>
-------------------------	---

REVIEWER	Dr Emma Lancashire Senior Research Fellow University of Birmingham UK
REVIEW RETURNED	24/02/2012

THE STUDY	<p>Appropriateness of statistical methods: I have ticked NO because I am not a statistician and do not therefore feel suitably qualified to respond YES.</p> <p>I ticked YES to the last point in my original review but on reading the question again I think I should have ticked NO as the way I read the question is that this is the response that means no action needs to be taken.</p>
GENERAL COMMENTS	With the exception of your response to point 2.17 I think you have

	<p>either made the suggested changes in my review or have provided appropriate justification for not doing so. With regard to 2.17 my comment was referring to the second 'At 18 months' in line 42 of page 13 of the original manuscript (or line 9 and 11 of the tracked changes version) and not the first as you had understood. I should have specified this but did not note at the time that there were 2 'at 18 months' in this line. I still think that for the 4th sentence of this paragraph (starting at the end of line 42 on page 13 of the original manuscript) it would be appropriate and clearer if it read: 'Compared to baseline, at 18 months the proportion of overweight ...'</p> <p>With regard to your response to point 2.7 I would like to clarify that I had not misunderstood the data presented in Table 3. The sentence 'When the anthropometric data are presented the change in mean value from baseline to follow up in each arm is presented which makes sense.' referred to page 13 lines 42 to 57 of the original manuscript (or p15 lines to 1 to 23 in the tracked changes version). I acknowledge that I should have made it clear that I was referring to the presentation of the results in the text and that also I misleadingly put 'mean value' when it should have been 'proportion' and I therefore think the your interpretation is understandable. In this section of the results the proportion overweight/obese and waist circumference > than or = to 85th percentile are compared in control and intervention group separately between baseline and the 2 points of follow up. I think presentation of the data in this way is helpful for the anthropometric data but can see that for the behavioural data the variables are not presented as proportion over a certain cut off and that it may not be appropriate to compare mean values for these variables for control / intervention groups between baseline and follow up. I accept that with such a small exploratory trial it was not possible to adjust for baseline values but are reassured that, should a definitive trial take place, it is the intention to make such adjustments.</p>
--	---

VERSION 2 – AUTHOR RESPONSE

Reviewer: Vera Araujo-Soares, PhD
Senior Lecturer in Health Psychology
Institute of Health & Society
Faculty of Medical Sciences
Newcastle University

No competing interests.

Happy with the way the authors dealt with the comments made to a previous version of this manuscript.

There are, nevertheless, some minor things that I would like to be considered by the authors.

1.1 In response to point 1.2 the authors provide more details “to the study design section in the methods (p8).” On page 8 it is stated: “The class teacher... daily reminders...”. For how long and how many reminders were sent out? Oral reminders; for the children to transmit to their parents? Or written reminders?

It has now been made clear that the reminders were oral and given daily during the 2 week period in which the opt out form needed to be returned (p.8).

More important: in response to point 1.5. the authors state that they now refer to the taxonomy. The authors should refer to and more crucially use the new taxonomy:

Michie, S., Ashford, S., Sniehotta, F. F., Dombrowski, S. U., Bishop, A., French, D. P. (2011). A refined taxonomy of behaviour change techniques to help people change their physical activity and healthy eating behaviours: the CALO-RE taxonomy. *Psychol Health* 26(11), 1479-1498
doi:10.1080/08870446.2010.540664.

It would also be key if the Behaviour Change Techniques listed on table 1 would map onto this taxonomy. It is not just a matter of referring to it, it is mostly a matter of using it for the sake of clarity within the description of the key BCT's within their intervention.

The new taxonomy has now been referenced with the older one (p.9) and the behaviour change techniques in table one now map on to those identified in the new taxonomy(p. 20).

Reviewer: Dr Emma Lancashire
Senior Research Fellow
University of Birmingham
UK

With the exception of your response to point 2.17 I think you have either made the suggested changes in my review or have provided appropriate justification for not doing so. With regard to 2.17 my comment was referring to the second 'At 18 months' in line 42 of page 13 of the original manuscript (or line 9 and 11 of the tracked changes version) and not the first as you had understood. I should have specified this but did not note at the time that there were 2 'at 18 months' in this line.

I still think that for the 4th sentence of this paragraph (starting at the end of line 42 on page 13 of the original manuscript) it would be appropriate and clearer if it read: 'Compared to baseline, at 18 months the proportion of overweight ...'

The wording has been changed to make it clearer (p.14)

With regard to your response to point 2.7 I would like to clarify that I had not misunderstood the data presented in Table 3. The sentence 'When the anthropometric data are presented the change in mean value from baseline to follow up in each arm is presented which makes sense.' referred to page 13 lines 42 to 57 of the original manuscript (or p15 lines 1 to 23 in the tracked changes version). I acknowledge that I should have made it clear that I was referring to the presentation of the results in the text and that also I misleadingly put 'mean value' when it should have been 'proportion' and I therefore think the your interpretation is understandable. In this section of the results the proportion overweight/obese and waist circumference > than or = to 85th percentile are compared in control and intervention group separately between baseline and the 2 points of follow up. I think presentation of the data in this way is helpful for the anthropometric data but can see that for the behavioural data the variables are not presented as proportion over a certain cut off and that it may not be appropriate to compare mean values for these variables for control / intervention groups between baseline and follow up. I accept that with such a small exploratory trial it was not possible to adjust for baseline values but are reassured that, should a definitive trial take place, it is the intention to make such adjustments.