PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	The Rhythm and Movement for Self-Regulation (RAMSR) intervention for preschool self-regulation development in disadvantaged communities: A clustered randomised controlled trial study protocol
AUTHORS	Williams, Kate E; Savage, Sally; Eager, Rebecca

VERSION 1 – REVIEW

REVIEWER	Catherine Gunzenhauser Freiburg University, Germany
REVIEW RETURNED	07-Feb-2020

GENERAL COMMENTS	Dear Dr. Williams and Team, I was honoured to be asked to review this promising study protocol. I think that your proposed research is relevant and has the potential for important practical implications. I also think that your design is generally appropriate to answer your research questions. However, I think there are a few issues that, to my mind, have not yet been described quite clearly (see below). I believe that clarifying these issues before the start of the research might help your project to succeed. Therefore, I have suggested minor revision. 1) Hypothesis C: you suggest that compared with teachers in the control group, teachers in the intevention group will show an "increase in teacher knowledge, confidence, practice, and attitudes related to self-regulation and rhythm and movement". Later on, you state that the control group will undergo a webinar on self-regulation development. However, self-regulation is just one of several suggested dependent variables. Thus, the topic of your "control intervention" does not address all your dependent variables in a balanced way. Related to this: It did not become quite clear to me why you think that teachers in the control group should undergo an active control condition (webinar) but children in the control group should, as far as I could see, not undergo an active control. 2) Power analysis: It is great that you provide power analysis for the sample size of children. However, as far is I understand, hypotheses C and D will include no analysis of child data but only analysis of teacher data. I could however find no power analysis for the sample size of teachers. I think it would be important to add this. 3) Power/Statistical analyses for Hypotheses A and B (primary and secondary outcomes of the intervention at the child level): You have many outcome variables, and in some cases one construct

	will be assessed using two instruments (e.g., behavioral regulation). While I think it is great that you are so comprehensive, I did not yet understand how exactly you will be including all outcomes in your model. It seems to me that your estimated sample size (N = 200) will not allow for having all outcomes in one model. Are you planning to run separate models for each outcome? Or will some construct be combined to indicators before paht analyses? 4) Statistical analysis for Hypothesis D (.."The proposed training and coaching support for teachers will result in high rates of intervention fidelity."). Here, it did not become clear to me how "high rates" will be defined. Will there be a specific cutoff? I hope that this is helpful and I wish you success with your project!

REVIEWER	serhat türkoğlu selcuk university medical faculty, Turkey
REVIEW RETURNED	08-Feb-2020

GENERAL COMMENTS	Although detailed information has been given in the article about the intended subjects, there is not enough information about how the evaluation has been made. Also, the discussion section for such a clear goal is quite insufficient. There are no comments on how it affects.

REVIEWER	James White Cardiff University, Wales, UK.
REVIEW RETURNED	15-Apr-2020

GENERAL COMMENTS	Abstract 1. Add what control group are receiving even if usual practice 2. Multiple primary outcomes are listed – please clarify if this results in one outcome measure or are these a number of measure which are all being treated as primary outcomes. The later requires a different approach to the analysis 3. Add ISCTRN registration number, assuming you have one. Introduction 4. Clarify how “self-regulation difficulties” are defined. 5. Page 5 lines 16-26: include more detail on the study designs for the studies described 6. Page 5 lines 47-51 – were these results from RCTs – please clarify 7. Expand and consider including a table of the pilot study results – show how these have informed the design of the study Method 8. Page 8 lines 31-32- please clarify here what strata are, or if yet to be defined, how will they be defined 9. Page 15 – lines 8-10: please clarify how strata of teacher numbers will be derived 10. Justify why you are stratifying on teacher enrolment numbers 11. Justify the need for the active control – is this usual practice in Australia? 12. Consider describing the intervention using the TIDIER framework: https://www.equator-network.org/reporting-guidelines/tidier/ 13. Page 15 lines 38-40. Consider whether you should have the same member of the research team coaching and doing fidelity assessments. Is this role conflicted? 14. Consider including an intervention logic model

	15. Provide dates when recruitment started and whether it is ongoing. 16. Sample size: a. Define “small intervention effect” b. Provide the parameters used to calculate the design effect: ICC, mean cluster size c. Provide coefficient of variation if used d. Clarify what power you are going for 80%/ 90% ? 17. Statistical analysis a. clarify how outcome measures are constructed and type of measure – continuous/ binary/ count etc. b. clarify what covariates will be adjusted for – strata used in randomisation? c. Do you intend to do a per protocol analysis? d. Clarify which models you will apply to which outcomes with SEM. For example are you using counts of binary outcomes.

REVIEWER	Muneera A. Rasheed Aga Khan University
REVIEW RETURNED	23-Apr-2020

GENERAL COMMENTS	The paper is clear and well-written overall covering all major areas. Some feedback to strengthen the paper is given below: Abstract: Under the introduction, it is not clear if the intervention in the current study is what was piloted and if it was effective? The manuscript indicates in this study the difference is the delivery staff. That can be made clearer. Similarly, it is confusing if the intervention content is being changed for the current study or was RAMSR delivered during pilot too? Has the trial been registered? If yes, please mention in the abstract. Teacher-reported outcomes may also be seen as a limitation and can be listed under limitations. If the forms the teachers are required to complete are beyond their routine assessments this may be a challenge too. Background A major switch in the current study from the pilot is training teachers keeping in view sustainability of the intervention. It would enhance the background if the authors expand on this aspect. Methods The Methods/Design can be renamed to be just Methods. Under Table 1 a bit more specificity about the intervention feasibility measures will be helpful though it is understandable that some of these issues cannot be decided pre-hand. Intervention Training: Is the training just one time or are any refreshers planned? Intervention delivery: The setting of the intervention delivery is not clear. Will it happen in the classrooms? At what time is it

	expected? Will children spend extra hours for the intervention or will the time-table be adjusted to include these intervention sessions? A lot of the measures depend on teachers who deliver the intervention but also new teachers at follow-up. While this is how the intervention will be scaled-up, this may pose a challenge to follow-up completion rates. I hope this has been considered and been discussed with the relevant authorities. Discussion I would expand the last paragraph of the discussion to talk more about scale-up challenges and opportunities with this intervention but also when implemented through teachers at work places. What can be some valuable lessons learnt for the implementation aspect and partnership with the schools? Intervention while may be low cost and be accessible may not be readily translated for scale-up due to operational feasibility. Barrier and facilitators to implementation can be part of the intervention implementation measures in Table 1.

VERSION 1 – AUTHOR RESPONSE

Reviewer comment		**Response**	
Reviewer 1			
3	Hypothesis C: you suggest that compared with teachers in the control group, teachers in the intervention group will show an "increase in teacher knowledge, confidence, practice, and attitudes related to self-regulation and rhythm and movement". Later on, you state that the control group will undergo a webinar on self-regulation development. However, self-regulation is just one of several suggested dependent variables. Thus, the topic of your "control intervention" does not address all your dependent variables in a balanced way. Related to this: It did not become quite clear to me why you think that teachers in the control group should undergo an active control condition (webinar) but children in the control group should, as far as I could see, not undergo an active control.	On consideration and in light of this comment and also Comment 18 below, as well as industry consultation since submission of the protocol, the active control component has now been removed.	

Reviewer comment		Response	
4	Power analysis: It is great that you provide power analysis for the sample size of children. However, as far as I understand, hypotheses C and D will include no analysis of child data but only analysis of teacher data. I could however find no power analysis for the sample size of teachers. I think it would be important to add this.	In relation to Hypothesis C we now explain under sample size: *With only up to eight teachers involved in the study, split across intervention and control, the study is underpowered to address hypothesis C using significance testing. Instead, we will use a match case series approach and descriptive statistics to take an exploratory approach to understanding any baseline similarities and differences in the teacher constructs among intervention teachers and matched control teachers, and change scores in constructs across baseline to follow-up.* For Hypothesis D, sample size is not relevant as we have now specified a specific cut point at which the hypothesis will be supported, thank you to your feedback at Point 6.	
5.	Power/Statistical analyses for Hypotheses A and B (primary and secondary outcomes of the intervention at the child level): You have many outcome variables, and in some cases one construct will be assessed using two instruments (e.g., behavioral regulation). While I think it is great that you are so comprehensive, I did not yet understand how exactly you will be including all outcomes in your model. It seems to me that your estimated sample size (N = 200) will not allow for having all outcomes in one model. Are you planning to run separate models for each outcome? Or will some construct be combined to indicators before path analyses?	Thank you for this important point. The new Table 3 takes the child outcomes measures that were previously a part of Table 2 and now provides additional details on how various component measures will be reduced to construct scores (see column 1). This will reduce the number of models required. These changes have been highlighted in the marked up version, rather than using tracked changes, for clarity.	
6.	Statistical analysis for Hypothesis D (.."The proposed training and coaching support for teachers will result in high rates of intervention fidelity."). Here, it did not become clear to me how "high rates" will be defined. Will there be a specific cut-off?	We have added a cut off here based on pilot findings, which has been detailed under Hypothesis D. *The proposed training and coaching support for teachers will result in high rates of intervention fidelity defined as ratings of 'implemented according to plan or with*	

Reviewer comment		Response	
		planned extensions' for 80% or more of the implementation instances for each activity across the program	
Reviewer 2			
7.	Although detailed information has been given in the article about the intended subjects, there is not enough information about how the evaluation has been made. Also, the discussion section for such a clear goal is quite insufficient. There are no comments on how it affects.	Apologies but we do not understand the specific recommended changes suggested here.	
Reviewer 3			
	Abstract		
8.	Add what control group are receiving even if usual practice	Thank you this has been added. <u>Control:</u> *usual practice kindergarten program.*	
9.	Multiple primary outcomes are listed – please clarify if this results in one outcome measure or are these a number of measure which are all being treated as primary outcomes. The later requires a different approach to the analysis	Thank you for this important point. Please see Comment 5 above for our response.	
10.	Add ISCTRN registration number, assuming you have one.	Trial registration details have been added to the abstract: *This trial is registered with the Australian New Zealand Clinical Trials Registry, ACTRN12619001342101.*	
	Introduction		
11	Clarify how "self-regulation difficulties" are defined.	Now added: *self-regulation difficulties including below average attentional and emotional regulation skills that do not improve as expected over time*	
12.	Page 5 lines 16-26: include more detail on the study designs for the studies described	All studies cited were RCTs and this has now been clarified briefly in the manuscript.	

Reviewer comment		Response	
13.	Page 5 lines 47-51 – were these results from RCTs – please clarify	We have provided further details of the initial published trial of the intervention: *Our first quasi-experimental study was a clustered trial in three preschool centres in disadvantaged communities, with two classrooms in each centre (22). One classroom in each centre was non-randomly assigned to intervention, and one to control, based on classroom schedules aligned with availability of visiting RAMSR specialists delivering the program.*	
14.	Expand and consider including a table of the pilot study results – show how these have informed the design of the study	While we have not included an additional table we now provide more detail on the RAMSR pilot findings and implications for the RCT as follows: *The RAMSR pilot was a quasi-experimental clustered trial in three kindergarten centres in disadvantaged communities, with two classrooms in each centre (22). One classroom in each centre was non-randomly assigned to intervention, and one to control, based on classroom schedules aligned with availability of visiting RAMSR specialists delivering the program. Results showed positive intervention effects for teacher-reported emotional regulation growth (moderate effect size of .35) and the directly assessed executive function of shifting for boys (large effect size of .60) (22). Fidelity ratings by visiting specialists delivering RAMSR showed that activities were implemented in accordance with the plan for 77% to 98% of activities instances, depending on the activity and child enjoyment was rates as high for 77% of all sessions conducted.* *This study will build on this previous trial by investigating the effectiveness of the same intervention as delivered by teachers, rather than visiting specialists. For activities with lower levels of fidelity across the pilot, these were adjusted and redesigned in preparation for this RCT, based on the qualitative notes provided by implementation leaders during the*	

Reviewer comment		Response	
		pilot. New evidence that will be produced by this RCT, is important to establish if the intervention is to be scaled up and increase its reach with the more sustainable approach of supporting existing teachers to deliver the program.	
	Method		
15.	Page 8 lines 31-32- please clarify here what strata are, or if yet to be defined, how will they be defined	Stratification of enrolment capacity (teachers work with either 22 or 44 children) has now been further explained in a number of locations to provide greater clarity. Please also see Comment 17 below. See Kindergarten teacher recruitment section.	
16.	Page 15 – lines 8-10: please clarify how strata of teacher numbers will be derived	Teachers will provide their enrolment capacity on the expression of interest form. This has now been clarified. See Kindergarten teacher recruitment section.	
17.	Justify why you are stratifying on teacher enrolment numbers	Thank you for this suggestion, the following has been added, and please see Comments 15 and 16 above, we hope this is now clearer. *This is because teachers will either teach 22 children or 44 children across a given week depending on their centre or work arrangements as described above. Stratification will mean an equal number of smaller units and larger units will be assigned to each condition to achieve a balanced design.*	
18.	Justify the need for the active control – is this usual practice in Australia?	In light of Comment 3 above, and your feedback, as well as industry consultation since submission of the protocol, the active control component has now been removed.	
19.	Consider describing the intervention using the TIDIER framework: https://www.equator-network.org/reporting-guidelines/tidier/	We carefully reviewed the TIDieR guidelines prior to submission and have covered all aspects in the text. We have now additionally added the url for further access to RAMSR training etc. *Further and regularly updated information on RAMSR, including training approaches and*	

Reviewer comment		Response	
		opportunities, can be found at https://ramsrblog.wordpress.com/	
20.	Page 15 lines 38-40. Consider whether you should have the same member of the research team coaching and doing fidelity assessments. Is this role conflicted?	We do not consider this role conflicted. The session plans that teachers follow are highly specific and ratings of the extent to which they are followed can be made in a relatively objectively and factual way. However, we acknowledge that an additional level of fidelity checking outside of the teacher-coach relationship might be valuable and so have added: *The research team lead will also make additional visits to all pairs of teachers and coaches to collect additional independent fidelity checklist information.*	
21.	Consider including an intervention logic model	Thank you for this suggestion. We have now added a logic model as Figure 1.	
22.	Provide dates when recruitment started and whether it is ongoing.	This has now been clarified. *Teachers will invite all children enrolled and commencing the 2020 kindergarten year between 28th January and the 21st of February 2020 (first four weeks of the kindergarten year) to participate in the research through gaining parental consent.*	
23.	Sample size: a. Define “small intervention effect” b. Provide the parameters used to calculate the design effect: ICC, mean cluster size c. Provide coefficient of variation if used d. Clarify what power you are going for 80%/ 90% ?	We have added additional detail to this section to address (we do not use a coefficient of variation). *Assuming a simple sampling design with no clustering, calculation in G*Power indicates a required sample size of 76 to detect small intervention effects (.14) with a power of 0.8 using the planned analytic approach (path analysis equivalent to multiple regression). The balanced cluster design of the RCT produces a design effect of up to 1.9 (using average ICC for same measures used in the pilot of 0.3 and average cluster size of 33), increasing the effective sample size to 144.*	

Reviewer comment		Response	
23.	Statistical analysis a. clarify how outcome measures are constructed and type of measure – continuous/ binary/ count etc. b. clarify what covariates will be adjusted for – strata used in randomisation? c. Do you intend to do a per protocol analysis? d. Clarify which models you will apply to which outcomes with SEM. For example are you using counts of binary outcomes?	a) and d) Please see the new Table 2 which details the child outcome measures and how various component measures will be reduced to construct scores (see column 1) and the nature of the outcome variable for each construct. This will reduce the number of models required. *b)* Now added to the document: *Covariates will be selected from child gender and age, level of family income, parental education, non-English speaking status, and Aboriginal status. Selection will be based on consistency and significance of bivariate correlations among these covariates and outcomes measures.* c) As noted the approach to be taken is intention-to-treat rather than per protocol.	
Reviewer 4			
	Abstract:		
24.	Under the introduction, it is not clear if the intervention in the current study is what was piloted and if it was effective? The manuscript indicates in this study the difference is the delivery staff. That can be made clearer.	This has now been made more clear in the abstract.	
25.	Similarly, it is confusing if the intervention content is being changed for the current study or was RAMSR delivered during pilot too?	This has now been made clearer, it is the same intervention.	
26.	Has the trial been registered? If yes, please mention in the abstract.	Yes, the trial has been registered and this information is now found at the end of the abstract.	
27.	Teacher-reported outcomes may also be seen as a limitation and can be listed under limitations.	An additional limitation has now been added related to this.	

Reviewer comment		Response	
28.	If the forms the teachers are required to complete are beyond their routine assessments this may be a challenge too.	The pilot project found that rates of data completion by teachers, completing forms for each child, was very high. We are not concerned about this.	
	Background		
29.	A major switch in the current study from the pilot is training teachers keeping in view sustainability of the intervention. It would enhance the background if the authors expand on this aspect.	*This aspect has now been expanded upon in the abstract and additional details has been added in the 7th paragraph of the Background, as follows:* *This study will build on this previous trial by investigating the effectiveness of the same intervention as delivered by teachers, rather than visiting specialists. For activities with lower levels of fidelity across the pilot, these were adjusted and redesigned in preparation for this RCT, based on the qualitative notes provided by implementation leaders during the pilot. New evidence that will be produced by this RCT, is important to establish if the intervention is to be scaled up and increase its reach with the more sustainable approach of supporting existing teachers to deliver the program. While training teachers to deliver interventions with fidelity has challenges (23), in this study we will use best evidence in teacher professional development (24) to ensure fidelity and quality of intervention implementation. Specifically the training takes an experiential model, followed up with supports including coaching and a community of practice. An implementation science approach will be used to understand the enablers and barriers to teachers implementing the intervention within their usual programming, and what is needed to support their confidence and skill to do so.*	
	Methods		
30.	The Methods/Design can be renamed to be just Methods.	Thank you this has now been changed.	

Reviewer comment		Response	
31.	Under Table 1 a bit more specificity about the intervention feasibility measures will be helpful though it is understandable that some of these issues cannot be decided pre-hand.	Additional information has been added: *Each of the 6 activities undertaken in each section will require a rating by teachers of either 'as per plan'; 'as per plan with adjustments'; 'not implemented'.*	
32.	Intervention Training: Is the training just one time or are any refreshers planned?	There is an additional brief refresher which has now been added. *Prior to the implementation period an online refresher course will be completed.*	
33.	Intervention delivery: The setting of the intervention delivery is not clear. Will it happen in the classrooms? At what time is it expected? Will children spend extra hours for the intervention or will the time-table be adjusted to include these intervention sessions?	Additional details have been added to clarify this: *Teachers will deliver this as part of their usual group floor or circle time within the kindergarten daily routine. Typically this is held in the morning as a warm-up to the day.*	
34.	A lot of the measures depend on teachers who deliver the intervention but also new teachers at follow-up. While this is how the intervention will be scaled-up, this may pose a challenge to follow-up completion rates. I hope this has been considered and been discussed with the relevant authorities.	Yes this has been considered and as the research team has strong engagement with both prior-to-school and school authorities who approve and support research we have the experience and support ensure adequate completion rates.	
	Discussion		
35.	I would expand the last paragraph of the discussion to talk more about scale-up challenges and opportunities with this intervention but also when implemented through teachers at work places. What can be some valuable lessons learnt for the implementation aspect and partnership with the schools? Intervention while may be low cost and be accessible may not be readily translated for scale-up due to operational feasibility. Barrier and facilitators to implementation can be part of the intervention implementation measures in Table 1.	Thank you for this important note. Indeed we do intend to investigate barriers and facilitators as part of the teacher narrative interviews and this has now been noted in Table 1 as you suggest. We have also added in the last paragraph: *Further, the study will contribute important information about the particular barriers and facilitators that early educators face in the implementation of specific intervention programs as they aim to scale up. Increased understanding of implementation factors will be of benefit to educational intervention designers that seek to embed enablers and*	

Reviewer comment		Response	
		address potential barriers as part of design and implementation processes.	

VERSION 2 – REVIEW

REVIEWER	Catherine Gunzenhauser Freiburg University, Germany
REVIEW RETURNED	03-Jun-2020

GENERAL COMMENTS	Thank you very much for the opportunity to review the revised version of this research protocol. In my opinion, the authors have done a good job addressing all my previous comments. I think this is a relevant and methodologically sound study. I have just one minor additional comment: I think it would be helpful if the authors could explain and discuss in a more explicit way why they do not use the same measures in the 8-weeks follow up and in the 12-months follow up. I believe there might be pracitical reasons (some measures are validated for kindergarteners but not for first-graders; direct assessments are more difficult and harder to organize in first grade). After addressing this issue, I recommend acceptance of the study protocol.

REVIEWER	Muneera Rasheed Aga Khan University
REVIEW RETURNED	08-May-2020

GENERAL COMMENTS	The comments have been adequately addressed.

VERSION 2 – AUTHOR RESPONSE

Thank you to the reviewers for their time. There remains one minor suggestion from Reviewer 1 as follows:

Reviewer comment: I think it would be helpful if the authors could explain and discuss in a more explicit way why they do not use the same measures in the 8-weeks follow up and in the 12-months follow up. I believe there might be pracitical reasons (some measures are validated for kindergarteners but not for first-graders; direct assessments are more difficult and harder to organize in first grade).

Author response: We have now added the following to explain this: "Teacher-report has been selected as the only data collection approach at the 12 month follow-up period, rather than a repeat of direct child assessments under taken at baseline and eight-week follow-up. This is because it is anticipated that study children will attend up to 50 different schools spread across a wide geographic region and it is not feasible within financial constraints of the study to undertake direct child assessments. However, as depicted in Table 3, four of the teacher-report measures are repeated across all three time points which will allow for some consistency in the analyses."