

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

TITLE (PROVISIONAL)	Human resources and curricula content for early child development implementation: multi-country mixed-methods evaluation
AUTHORS	Kohli-Lynch, Maya; Ponce Hardy, Victoria; Bernal Salazar, Raquel; Bhopal, Sunil; Brentani, Alexandra; Cavallera, Vanessa; Goh, Esther; Hamadani, Jena; Hughes, Rob; Manji, Karim; Milner, Kate; Radner, James; Sharma, Sonia; Silver, Karlee; Lawn, Joy; Tann, Cally J

VERSION 1 - REVIEW

REVIEWER	Amanda Thrift Professor, Monash University, Melbourne, Australia
REVIEW RETURNED	07-Jul-2019

GENERAL COMMENTS	<p>The authors detail an evaluation of the Savings Brain projects funded by Grand Challenges Canada. The objective was to specifically focus on human resources and content of the curricula for early childhood development. There are few such large evaluations of projects being scaled up, and the findings add to what is currently known in the literature. Linkage between the framework and evaluation is unclear, and such a linkage would add to the accessibility and value of the manuscript. The following is meant to help improve the quality of the manuscript:</p> <ol style="list-style-type: none">1. Web appendix Figures A and B would benefit from a legend to explain the Framework and Theory of Change, including definitions of the abbreviations used. All abbreviations in all tables and web tables should also be defined.2. The manuscript would benefit from linkage between the Framework, Theory of Change, and Results. The Theory of Change figure includes "Metric Numbers" that seem to have no connection to the evaluation being conducted. Perhaps the authors could include numbers that then relate to the Framework and to each element of the results, so that readers can better determine where each element of the evaluation contributes to the frameworks. That is, how does each component relate to the Tables, figures and text?3. Web Appendix Figure 2. It is currently difficult to determine which arrows correspond to the 'flow of learning' and which correspond to 'transition to scale' as the arrows are similar in colour. I would suggest modifying one of the colours so that the difference is clear.4. Web Appendix Table A would benefit from an extra column that provides details on the information used for this manuscript. That is, how many service delivery forms were used across all projects? How many focus groups were conducted and what was the average
-------------------------	---

	<p>number of people per focus group? This would provide a useful summary for readers.</p> <p>5. Given that the results of the themes and subthemes arising from the quantitative and qualitative analyses (Table 2) were used to obtain quantitative data from the Transition-to-scale projects, it might be better to swap around Tables 1 and 2. It would be helpful to also highlight the connection between the two tables.</p> <p>6. Currently it is unclear where the projects were conducted (except that we know they were conducted in 17 LMICs). It would be useful to include a table summarising the projects in the Saving Brains program, e.g. country, type of study (seed vs transition-to-scale), etc.</p> <p>7. Figure 2c. In the 'contribution to expenses' group of projects, how were there more people who completed training than there were who started training? Or is this an error?</p>
--	--

REVIEWER	Yeon-Gyu, Jeong Department of Physical Therapy, Yeosu Institute of Technology, Republic of Korea.
REVIEW RETURNED	08-Jul-2019

GENERAL COMMENTS	I think that the manuscript is clearly written for the purpose that you argue. And the proper methods were provided. However, I want for all readers to understand your manuscript, because the several sentences were vague and difficult to understand for all readers.
-------------------------	---

REVIEWER	Eileen Stuart-Shor University of Massachusetts Boston USA
REVIEW RETURNED	03-Oct-2019

GENERAL COMMENTS	<p>Human resources and curricula content for early child development implementation: multi-country mixed-methods evaluation.</p> <p>This is an important and timely topic and I applaud the authors for attempting to shed light on the complexities of implementing early child development programmes in low and middle income countries. The authors provide a robust descriptive analysis of systems and process level factors which influence implementation but I think the paper, by virtue of some of the wording, implies inferential results and impact which I think goes beyond the study design and analysis. The discussion section of the paper needs to be strengthened in order to understand the significance of their findings. The large sample size and the rigor of quantitative and qualitative analysis are strengths of the study. The findings are very helpful but I think it needs to be more clear that these are descriptive findings and that one can't make inferences from the data in terms of which factors are linked to impact or effectiveness.</p> <p>In the background the authors note that there is "limited data to guide practical implementation" and there is a need for "guidance for contextual adaptation". I'm concerned that the wide variances in process and program factors and the fact that the study does not evaluate impact or effectiveness, make it difficult to say that they achieved these objectives. I think the data they have is important, and helpful in a preliminary way, but I think the findings beg for more</p>
-------------------------	---

studies that actually look at differences in outcomes based on workforce (professional cadre, training) and curricula (design, duration, dose) factors.

The authors acknowledge the limitation of heterogeneity in program design and curricula in the paper, but I think they need to consider this more in the way they write up the results and discussion.

Suggestions:

- In the abstract the statement in the results section about the major themes for the curricula content needs more context. The study was not designed in a way that allows you to see if flexible adaptation of content and delivery, maintaining fidelity and intervention duration and dosage makes a difference in terms of outcome or impact. Perhaps in the methods section or last section of the background you could say that the purpose of the study was to describe the human resource and curricula content of these programs. Somehow it needs to be clearer that the themes you found are not related to outcomes/impact but that you are reporting on the ways this has been implemented across 32 programs in 4 countries. To me it's a strong statement (and a bit of over reach of the data) in the conclusion that the mixed-method evaluation of a multi-country ECD portfolio identified themes for policymakers and that can guide implementation

- The manuscript, particularly the background but throughout the entire document has too many abbreviations which makes it hard to follow your logic.

- In the background and methods I would have also found it helpful to understand a bit more about the Saving Brains portfolio. I found the Web appendices on their conceptual model and theory of change extremely helpful; perhaps a very brief overview or introduction to this in the text in the background section could be added.

- P9/10; the percent (35%, 66%) add up to 101%. This is likely due to a round error but it needs to add up to 100%. The same issue came up in other places where you presented a breakdown by percent (e.g. p 14; 61%, 20%, 20%); please review and correct.

- The quantitative and qualitative methods (data sources, data management and analytic plan) are well described. In the qualitative methods it would be helpful to say if member-checking was used as a way to assure the credibility and trustworthiness of the analysis.

- Ethics; since this work was done in LMIC it is important to show that ethics review was done in the LMICs. This may be implied in the 2nd sentence, but it should be explicitly stated.

- The degree in variance described in Table 1 in terms of human resources and curricula content is what makes it hard to understand the significance of the study outcomes.

- The title implied to me that I would gain important insights from this multi-country study that would inform the program I could implement in my LMIC. While the manuscript did an excellent job of descriptive analysis of the data they had; design issues make it hard to think about the type of program (curriculum and program design that is most effective/practical to implement). It provides important insights in workforce/human resource issues that affect programs, but these themes are also impacted by the lack of information on the specific program's design. For instance, if you get the same results from 25 sessions of 50 minutes each as you get from 2 sessions of 60 minutes this has huge implications for the number of staff you need and for sustainability and impact how you might interpret some of the qualitative findings you have.

- P 18/19. Last sentence. You write that workforce decisions

	<p>around delivery of RCEL programmes can have substantial bearings on programme sustainability and impact. I agree with that statement but I do not think your design allows you to make comments on impact. It may be the wording that is the problem. I think you can say that your findings allow the reader to see how this has been implemented in various settings; and the perceptions of people in those programs around workforce and curricula; but beyond describing the program and perceptions I do not think you can make further conclusions/recommendations.</p> <ul style="list-style-type: none"> • P. 21 paragraph 2; line 2: Perhaps in your recommendations you can suggest research that focuses on the same endpoints/outcomes that they used in these studies that showed effectiveness. • What your finding do clearly state is the need to do further research to understand what human resource (cadre, training) and curricula (desgn, duration, dose) have the best outcomes/impact. • I think the conclusions should start off with acknowledging that this is descriptive data and the hetergenity of the programs makes cross-program comparisons difficult; but common themes that emerged were (what you found). The last sentence seems to me to be an overreach based on the data you have.
--	---

VERSION 1 – AUTHOR RESPONSE

Author responses to Reviewer 1 comments:

Comment 1.1: Web appendix Figures A and B would benefit from a legend to explain the Framework and Theory of Change, including definitions of the abbreviations used. All abbreviations in all tables and web tables should also be defined. Thank you for the suggestion. Legends have now been added to ensure all abbreviations and definitions are clear throughout the figures and tables.

Response 1.1: Thank you for the suggestion. Legends have now been added to ensure all abbreviations and definitions are clear throughout the figures and tables.

Comment 1.2: The manuscript would benefit from linkage between the Framework, Theory of Change, and Results. The Theory of Change figure includes “Metric Numbers” that seem to have no connection to the evaluation being conducted. Perhaps the authors could include numbers that then relate to the Framework and to each element of the results, so that readers can better determine where each element of the evaluation contributes to the frameworks. That is, how does each component relate to the Tables, figures and text?

Response 1.2: Whilst we appreciate this thoughtful comment, unfortunately it is not as straightforward as mapping the numbers to the manuscript. We have added a sentence clarifying that the Theory of Change ‘metric numbers’ map only to the RMAF data and, as not all results come from RMAF data (and in fact a large proportion do not), stating a ‘metric number’ for each result would be somewhat misleading. Source of data is however stated beneath each figure and table.

Comment 1.3: Web Appendix Figure 2. It is currently difficult to determine which arrows correspond to the 'flow of learning' and which correspond to 'transition to scale' as the arrows are similar in colour. I would suggest modifying one of the colours so that the difference is clear.

Response 1.3: The Theory of Change was developed and designed by the Saving Brains platform team and not by this paper's authors, therefore unfortunately we cannot change the colours of the arrows. A footnote has been added to this figure to clarify this.

Comment 1.4: Web Appendix Table A would benefit from an extra column that provides details on the information used for this manuscript. That is, how many service delivery forms were used across all projects? How many focus groups were conducted and what was the average number of people per focus group? This would provide a useful summary for readers.

Response 1.4: Thank you for the suggestion. A column has been added to Web Appendix Table A providing details on data sources.

Comment 1.5: Given that the results of the themes and subthemes arising from the quantitative and qualitative analyses (Table 2) were used to obtain quantitative data from the Transition-to-scale projects, it might be better to swap around Tables 1 and 2. It would be helpful to also highlight the connection between the two tables.

Response 1.5: Thank you for the suggestion. We agree that is a better order in which to introduce the tables. We have reordered the tables and edited references to them in the text, including the connection between the two tables.

Comment 1.6: Currently it is unclear where the projects were conducted (except that we know they were conducted in 17 LMICs). It would be useful to include a table summarising the projects in the Saving Brains program, e.g. country, type of study (seed vs transition-to-scale), etc

Response 1.6: Thank you for the suggestion. A summary table of the Saving Brains projects has been published in the supplementary material of 'Milner K, Bernal R, Brentani A, Britto P, Dua T, Gladstone M, et al. Contextual design choices and partnerships for scaling early child development programmes. *Arch Dis Child*. 2019;104(Suppl 1):S3-12'. This has now been referenced in the methods section of this paper. Additionally, a map of countries where projects were implemented has been added (now Figure 2).

Comment 1.7: Figure 2c. In the 'contribution to expenses' group of projects, how were there more people who completed training than there were who started training? Or is this an error?

Response 1.7: Thank you for this observation which we think is in reference to original Figure 2d (now Figure 3d). Having gone back to the data, it appears that one project reported 26 'entered', 40 'completed' and 32 'delivered'. As these numbers do not appear to be correct, these data have now been excluded from Figure 3d.

Author responses to Reviewer 2 comments:

Comment 2.1: I think that the manuscript is clearly written for the purpose that you argue. And the proper methods were provided. However, I want for all readers to understand your manuscript, because the several sentences were vague and difficult to understand for all readers.

Response 2.1: Thanks. We have reviewed the manuscript and have amended where necessary for clarity.

Author responses to Reviewer 3 comments:

Comment 3.1: In the abstract the statement in the results section about the major themes for the curricula content needs more context. The study was not designed in a way that allows you to see if flexible adaptation of content and delivery, maintaining fidelity and intervention duration and dosage makes a difference in terms of outcome or impact. Perhaps in the methods section or last section of the background you could say that the purpose of the study was to describe the human resource and curricula content of these programs. Somehow it needs to be clearer that the themes you found are not related to outcomes/impact but that you are reporting on the ways this has been implemented across 32 programs in 4 countries. To me it's a strong statement (and a bit of overreach of the data) in the conclusion that the mixed-method evaluation of a multi-country ECD portfolio identified themes for policymakers and that can guide implementation

Response 3.1: Thanks for this helpful comment. It is true that impact/outcome data was not available for the majority of projects at the time of the evaluation and it was not the part of the research objectives of this work. A sentence has now been added to the end of the first paragraph clarifying this more explicitly. We accept the comment around the relevance to policymakers 'to guide implementation' but do feel that the data presented provide information useful for policymakers more broadly. In response, in the abstract conclusion we have softened the wording to say that we identified themes 'for consideration by policymakers and programmers'. We hope this clarifies that our findings do not provide recommendations on human resource/curricula issues. We have also added a sentence on aims to the background section of the abstract to clarify that our findings do not relate to impact.

Comment 3.2: The manuscript, particularly the background but throughout the entire document has too many abbreviations which makes it hard to follow your logic.

Response 3.2: In response to this helpful observation we have tried to rationalise these, including removing SDG, WHO, CHW. Others e.g. RCEL, ECD, GCC are necessary as they recur many times in the paper and are quite long so would take up word count if not abbreviated.

Comment 3.3: In the background and methods I would have also found it helpful to understand a bit more about the Saving Brains portfolio. I found the Web appendices on their conceptual model and theory of change extremely helpful; perhaps a very brief overview or introduction to this in the text in the background section could be added.

Response 3.3: We are pleased that the reviewer found the web appendices helpful and appreciate their request for more information on the Saving Brains portfolio more widely. In response, we have added a short sentence at the beginning of the methods section under 'overview of saving brains portfolio evaluation' and referenced Radner et al paper which describes the portfolio in more depth, and added a reference to a summary table of all Saving Brains projects (contained in Milner K, Bernal R, Brentani A, Britto P, Dua T, Gladstone M, et al. Contextual design choices and partnerships for scaling early child development programmes. Arch Dis Child. 2019;104(Suppl 1):S3-12). We hope that these are helpful and provide the necessary information.

Comment 3.4: P9/10; the percent (35%, 66%) add up to 101%. This is likely due to a round error but it needs to add up to 100%. The same issue came up in other places where you presented a breakdown by percent (e.g. p 14; 61%, 20%, 20%); please review and correct.

Response 3.4: Thank you and apologies for this oversight which, you are correct, is due to rounding. This has now been amended on p9/10. On p14, 61% is in reference to all workers, and the 20%s are in reference to lay community member workers only – we have now clarified this in text.

Comment 3.5: The quantitative and qualitative methods (data sources, data management and analytic plan) are well described. In the qualitative methods it would be helpful to say if member-checking was used as a way to assure the credibility and trustworthiness of the analysis.

Response 3.5: Thank you for the comment. While formal member-checking was not undertaken, interviewers and focus group facilitators summarised and reflected throughout data collection to improve validity of results. This has now been added to the methods section.

Comment 3.6: Ethics; since this work was done in LMIC it is important to show that ethics review was done in the LMICs. This may be implied in the 2nd sentence, but it should be explicitly stated.

Response 3.6: Thank you for the comment. We received approval from London School of Hygiene and Tropical Medicine ethics committee to use deidentified and anonymised data. We have removed

the second sentence in the ethics section to avoid misleading readers regarding ethics approval sought for this study.

Comment 3.7: The degree in variance described in Table 1 in terms of human resources and curricula content is what makes it hard to understand the significance of the study outcomes

Response 3.7: (Now Table 2) Yes, agreed, we have added this to the limitations section of the discussion, and to conclusion.

Comment 3.8: The title implied to me that I would gain important insights from this multi-country study that would inform the program I could implement in my LMIC. While the manuscript did an excellent job of descriptive analysis of the data they had; design issues make it hard to think about the type of program (curriculum and program design that is most effective/practical to implement). It provides important insights in workforce/human resource issues that affect programs, but these themes are also impacted by the lack of information on the specific program's design. For instance, if you get the same results from 25 sessions of 50 minutes each as you get from 2 sessions of 60 minutes this has huge implications for the number of staff you need and for sustainability and impact how you might interpret some of the qualitative findings you have.

Response 3.8: Thank you for this comment. We agree that it is difficult to apply the findings in this paper to the implementation of individual programmes. Our aim was to describe a diverse multi-country portfolio of RCEL projects, and due to that diversity, it is difficult to draw clear conclusions re human resources and curricula content for individual project implementation. We hope that this is made clearer from edits to limitations and conclusions.

Comment 3.9: P 18/19. Last sentence. You write that workforce decisions around delivery of RCEL programmes can have substantial bearings on programme sustainability and impact. I agree with that statement but I do not think your design allows you to make comments on impact. It may be the wording that is the problem. I think you can say that your findings allow the reader to see how this has been implemented in various settings; and the perceptions of people in those programs around workforce and curricula; but beyond describing the program and perceptions I do not think you can make further conclusions/recommendations.

Response 3.9: Thanks. We are saying here that Radner et al highlighted that workforce decisions have substantial bearings on sustainability and impact. We have reworded to make this clearer. We have also added a sentence at the end of the first paragraph of the methods to clearly communicate that it was not within the remit of the evaluation to look at outcomes or impact.

Comment 3.10: P. 21 paragraph 2; line 2: Perhaps in your recommendations you can suggest research that focuses on the same endpoints/outcomes that they used in these studies that showed effectiveness.

Response 3.10: Thanks for the suggestion. We have added a sentence recommending alignment of outcomes in future research.

Comment 3.11: What your finding do clearly state is the need to do further research to understand what human resource (cadre, training) and curricula (desgn, duration, dose) have the best outcomes/impact.

Response 3.11: Thanks. We have added this to the end of the conclusions.

Comment 3.12: I think the conclusions should start off with acknowledging that this is descriptive data and the hetergenity of the programs makes cross-program comparisons difficult; but common themes that emerged were (what you found). The last sentence seems to me to be an overreach based on the data you have.

Response 3.12: Thank you. We have added this to the beginning of the conclusion and removed the last sentence of the conclusion.

VERSION 2 – REVIEW

REVIEWER	Professor Amanda Thrift Monash University, Australia
REVIEW RETURNED	24-Jan-2020

GENERAL COMMENTS	<p>The authors detail an evaluation of the Savings Brain projects funded by Grand Challenges Canada. There are few such large evaluations of projects being scaled up, and the findings add to what is currently known in the literature. The authors have addressed many, but not all, of my initial concerns:</p> <ol style="list-style-type: none">1. Tables and Web appendices. There are still abbreviations in the tables and web tables that have not been defined (e.g. check title of Table 1). Please amend.2. Some attempt at clarifying how the framework contributed to the evaluation process would be useful to readers.
-------------------------	--

VERSION 2 – AUTHOR RESPONSE

Reviewer 1 Comments:

1 Tables and Web appendices. There are still abbreviations in the tables and web tables that have not been defined (e.g. check title of Table 1). Please amend.

Response from authors:

Thank you for this comment. We have now defined all abbreviations that appear in tables and web appendices.

2 Some attempt at clarifying how the framework contributed to the evaluation process would be useful to readers.

Response from authors:

Thank you for this suggestion. We have added elaborated in the 'Methods> Overview of the Saving Brains Portfolio evaluation' section to clarify how the conceptual evaluation framework contributed to the portfolio-level evaluation.