Reviewer’s report

Title: Depressive Symptoms and Negative Life Events: What Psycho-social Factors Protect or Harm Left-behind Children in China?

Version: 0 Date: 23 May 2017

Reviewer: Kirsten Hancock

Reviewer’s report:

General comments:

I found the manuscript to be an interesting account of left-behind children in China. In my view, the manuscript requires major revision before it is of an acceptable standard for publication in this journal or elsewhere. There are many examples of poor language or grammar throughout the manuscript, and extensive editing is needed. I don't have the resources to list them all here, perhaps the authors may consider using an English language editing service.

Language errors aside, I enjoyed reading the manuscript concerning depressive symptoms among Left Behind Children in China. Given the sheer number of children experiencing such circumstances throughout rural China, it's a worthy research topic. The authors also extend previous research by examining links between stressful life events and depression for left-behind and non-left-behind children, and examine different types of left-behind families.

A major concern is that almost half the sample is excluded because of incomplete information. No data is provided to let readers understand if this is a problem. Do you have any data on the non-completes? What are their characteristics? Are they more likely to be LBC? Are they children more likely to struggle with literacy and therefore will struggle completing multiple questionnaires?

The manuscript would greatly benefit by giving some attention to the magnitude of the effects reported (i.e. effect size, or is the effect meaningful?). Readers can always scale differences in outcomes back to a proportion of a standard deviation, but it's more helpful to do this for the reader. What does a 1-point difference mean? Is this a small or large effect? You have large sample sizes in most analyses, so it's not surprising that you find p-values less than .05, the real interest is in what it means.

Specific comments:

Abstract.

1. In the background section, there's no clear link between the first and second sentence. The objective of the study would be much clearer if you stated something like "With restricted family support, left-behind children in China are at greater risk of developing depressive symptoms. The
objective of this study was to explore how the relationship between stressful life events and depressive symptoms varies for different groups of left-behind children and non-left behind children.

2. The results section (in the abstract and major section) may be easier to follow if you describe results in terms of risk factors and protective factors, rather than X is predicted by Y.

3. Your conclusion that stressful life events is a crucial risk factor for depressive symptoms is not the most interesting aspect of the study - such a pattern applies to all children, not just those left-behind.

Introduction

1. The literature review is generally comprehensive in terms of breadth, but the quality of writing and grammar can be substantially improved. Many references are cited, but not adequately described in some places, so their contribution to the literature review is superficial. A few examples are given below, but the Background section should be revised throughout.

2. Initially, (in the abstract and the first page) I was confused about whether the authors were referring to being left behind as a stressful event (which it likely is), or if the interest was in how LBC respond to stressful events given that their parents have more limited opportunities to provide support when those stressful events occur. Reading the rest of the manuscript it became clear, but the distinction should be made up front.

3. Page 6, line 121. An example of better writing would be to say what the increased risk is, not just that it is higher. It's less interesting to me how many participants were in the sample, I'd rather read "He et al. reported that the likelihood of depression among students in grades 4-6 was two-three times higher among LBC than non-left behind children. More information is conveyed in fewer words.

4. Page 6, line 123 - does the Wang study report prevalence among non-LBC? I don't know how important the 23-24% is without reference to a comparison population.

5. Page 6, line 125. The authors state that 'depression is also related to social support etc…'. These statements need to be more specific - how are they related exactly?

6. Several sentences I didn't understand, e.g. line 156-159, 164-167, 232-233.

7. Lines 169-174 are repetitive and can be summarised in a single sentence.

8. Lines 188-192. I don't know that the Raja study about attachment to peers is an appropriate line of research to cite here. There is an extensive international literature about the positive effects of grandparent involvement in the lives of grandchildren - particularly in single-parent families, which would be a more relevant literature given your statement about children being looked after by family members.
9. Lines 216-227. This section is hard to follow. The hypotheses would be easier to follow if they were grouped into related questions. For example, hypothesis 1 and 2 could be combined to say 'We hypothesised that in line with previous research LBC would have more depressive symptoms than NLBC, and that the extent of the difference would vary by the type of parent migration.

10. I think in relation to point 9 above, the hypotheses are also confusing because the authors refer to 'the impacts of stressful life events' in H3 and H4. Initially this reads as though it's the impact of stressful life events on depressive symptoms, but it's not until the methods that the reader understands that the measure of life stress events includes the perceived impact of the event in general (not on depression).

11. The authors vary between 'depressive symptoms' and 'depressive levels' (e.g. lines 216 and 217). Use consistent terms.

Methods

1. Line 238-243. Is experimenters the right term? This makes it sound like children participated in experiments. Perhaps they could be referred to as survey administrators?

2. Line 269. Roughly half the sample is excluded because students did not complete all questionnaires. This is a major concern of the study. Do you have any demographic / migration data on those with complete and incomplete information?

3. How does the 67% of the sample being LBC compare to census data or similar rural area? This seems very high.

4. Line 289 - what do scores of 0,1,2 represent? E.g. never, sometimes, often?

5. Line 292 - these item numbers don't mean anything. Just say half the items were positively worded and reverse scored.


7. Line 304 - what do the 0-5 numbers represent?


9. Line 313. Why did you transform a continuous variable into a categorical variable? Some justification is required.

10. Lines 320 and 322 conflict.
Results

1. A general comment for the results section is about the meaning or magnitude of results reported. E.g. In table 2, the mean depression score for NLBC is 13.6 and for LBCB 14.5. This is a difference of 0.9. Is this a meaningful difference? It's about 0.1 of a SD, which translates to a small effect size.

2. Some analyses refer to household income, others refer to socioeconomic status as measured by family income. Socioeconomic status is not just income. Just call it income, not socioeconomic status.

3. "Most of the participants lived in countryside with married parents and siblings" is confusing. This makes it sound like students are living with married parents, yet the point of the study is that parents live elsewhere. Do you mean "Most participants had married parents and at least one sibling".

4. Table 1 - Do you have information on where/ who the 45% of LBCB who do not lodge at school live with? This would be very important in the context of this study, i.e. the availability of adult family members to help support the child.

5. It's not clear that table 2 are the results of an ANOVA as described, or would be better presented as a MANCOVA given the multiple dependent variables and that the authors adjust for covariates.

6. The language of 'impacts of SLE' makes the results hard to understand, as earlier comment indicated. It might be better to call it a total stress score. 'Impact' makes reader look for associations between variables.

7. The hypotheses relating to how relationships between stress and depressive symptoms, for example, vary by migration type would suggest that moderator analyses or interactions would be included in the results. Instead, analyses have been stratified by migration type. Stratifying analyses in this way does not answer the question of whether an association is larger for one group than another. The regression analyses should include main effects and interaction terms for migration type.

8. Results reported in 455-475 do not make sense and should be re-written.

9. Table 6 - why is the F-value reported instead of specific interaction effects?

Discussion

1. Line 500 - describing this finding as watershed overstates it a little. It's well known internationally that depression is more common in adolescence. The cut point of 12 years for this study was arbitrary, really, these findings about demographic patterns are not interesting or new and don't need to be discussed in great detail.
2. Line 504 - the authors didn't measure father love, but father involvement. These are not the same constructs.

3. Line 512-517 doesn't make sense.

4. Line 516. Why is having more than one child more difficult, explain.

5. The authors use the language of 'the predicting effect of X on Y' etc throughout the manuscript. It's a correlational study, and as the authors note, this does not imply causation. Better language would be 'the association between X and Y…'

6. 556-557. Parents are important, but plenty of children do well without their biological parents because they have other significant adults in their lives.

7. 585-591 - doesn't make sense, hard to follow.

8. Some discussions should be included about reverse causation - that depressed individuals may perceive a life stress event to be more impactful than non-depressed individuals.

**Are the methods appropriate and well described?**
If not, please specify what is required in your comments to the authors.

No

**Does the work include the necessary controls?**
If not, please specify which controls are required in your comments to the authors.

Yes

**Are the conclusions drawn adequately supported by the data shown?**
If not, please explain in your comments to the authors.

Yes

**Are you able to assess any statistics in the manuscript or would you recommend an additional statistical review?**
If an additional statistical review is recommended, please specify what aspects require further assessment in your comments to the editors.

I am able to assess the statistics

**Quality of written English**
Please indicate the quality of language in the manuscript:

Not suitable for publication unless extensively edited


Declaration of competing interests
Please complete a declaration of competing interests, considering the following questions:

1. Have you in the past five years received reimbursements, fees, funding, or salary from an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

2. Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

3. Do you hold or are you currently applying for any patents relating to the content of the manuscript?

4. Have you received reimbursements, fees, funding, or salary from an organization that holds or has applied for patents relating to the content of the manuscript?

5. Do you have any other financial competing interests?

6. Do you have any non-financial competing interests in relation to this paper?

If you can answer no to all of the above, write 'I declare that I have no competing interests' below. If your reply is yes to any, please give details below.

I declare that I have no competing interests.

I agree to the open peer review policy of the journal. I understand that my name will be included on my report to the authors and, if the manuscript is accepted for publication, my named report including any attachments I upload will be posted on the website along with the authors' responses. I agree for my report to be made available under an Open Access Creative Commons CC-BY license (http://creativecommons.org/licenses/by/4.0/). I understand that any comments which I do not wish to be included in my named report can be included as confidential comments to the editors, which will not be published.

I agree to the open peer review policy of the journal