Reviewer’s report

Title: Trends in childhood and adolescent internalizing symptoms: results from Swedish population based twin cohorts

Version: 1 Date: 20 May 2019

Reviewer: Bryan Rodgers

Reviewer's report:

I will restrict my comments to areas that I raised in my original review and to the responses made by the authors in their Response to Reviewers section. I focus entirely on the original points 2 and 4, as the remaining issues are of little relevance unless these two important points are addressed appropriately.

2) What is meant by the term "internalizing symptoms" is fundamental to the present manuscript and the (revised) paper does not convey a clear picture. More specifically, the extent to which parent reports are valid indicators of "depression and anxiety" in children is an issue that has clouded inquiry in this field for decades. With the benefit of so much past experience, one would hope that contemporary research does not fall into the same traps as in the past. Analysis of trends over time in internalizing symptoms requires some thought about what this concept embraces (and what it does not include). For the present investigation, only parent reports are available for children aged 9 years whereas parent and self-reports are available for age 15 years, meaning that the choice of informant is crucial to the interpretation of findings. The authors have failed to grasp this fundamental point and appear very selective in the literature they choose to cite and not to cite. For example, the defence of the citation of Hariz et al. (2013) as an indication of the validity of SCARED parent reports in indicating children with clinical disorders ignores what is written in that paper, specifically that the high correlation between children self-reports and parent ratings found for the Arabic version was not found previously for the English version of the scared (Birmaher et al., 1999). I agree with the present authors that the literature on this topic is complicated but their own report glosses over such complications and, as such, presents what is a potentially misleading picture. I am concerned that the authors state in their response "The aim of the study was to investigate any secular changes between cohorts, and therefore further validation of the scales used, were outside the scope of this paper." Given that the manuscript already uses citations to support the validity of the scales used in the study, what particular reason is there for NOT reporting the correlations between parent reports and self-reports at age 15?

I have also realised that the Sharp et al. (2006) reference, now numbered 25, is inappropriate in that this study only collected self-report data using the sMFQ but it has been cited in support of the validity of "parent-rated scores". Apologies for not pointing this out in my original review.
4) The authors reject my suggestion to report interaction terms between time and sex on the basis that their separate analyses of males and females were "sensitivity analyses". This is simply incorrect. They state in their response "The aim of the study was not to explore sex differences" but it would be difficult to find a more inaccurate phrase to describe this study. A quick glance at the Results and Conclusions of the Abstract undermines this statement. The reported separate analyses for boys and girls are NOT sensitivity analyses by any conventional use of that term. The authors appear not to understand that a single analysis including an interaction term can provide both a test of different slopes for males and females and also the results they already choose to report from separate analyses. First, this avoids the increased risk of Type 1 errors arising from multiple tests of significance. More importantly, it provides extra information not obtained by the separate analyses. For example, if a finding was obtained showing a greater slope in females compared with males this could help explain why earlier studies (possibly with less statistical power) have reported significant trends for girls but not for boys. The present paper concludes that "The results demonstrated some trends in internalizing symptoms across birth cohorts, for both boys and girls." To omit from this whether these trends differ (or not significantly so) between boys and girls is an incredible waste of information. To hammer home this point (in case it is not already obvious) the regression results reported in Table 2 show a beta of 0.054 for boys' SDQ self-reports and a value of 0.155 for girls, which is almost a three-fold difference. Are readers meant to ignore this difference because both are statistically significant at a particular P level? I would hope that the majority of readers are not so naïve.

**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.

No

**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:

Acceptable

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