Author's response to reviews

Title: Measurement invariance of the depressive symptoms scale during adolescence

Authors:

Jennifer Brunet (jennifer.brunet@uottawa.ca)
Catherine M Sabiston (catherine.sabiston@utoronto.ca)
Michael Chaiton (michael.chaiton@utoronto.ca)
Nancy CP Low (nancy.low@mcgill.ca)
Gisèle Contreras (gisele.contreras@umontreal.ca)
Tracie A Barnett (tracie.barnett@gmail.com)
Jennifer L O'Loughlin (jennifer.oloughlin@umontreal.ca)

Version: 4 Date: 20 February 2014

Author's response to reviews: see over
Author's response to reviews

**Title:** Measurement invariance of the depressive symptoms scale during adolescence

**Authors:**

Jennifer Brunet (jennifer.brunet@uottawa.ca)  
Catherine M Sabiston (catherine.sabiston@utoronto.ca)  
Michael Chaiton (michael.chaiton@utoronto.ca)  
Nancy CP Low (nancy.low@mcgill.ca)  
Gisèle Contreras (gisele.contreras@umontreal.ca)  
Tracie A Barnett (tracie.barnett@gmail.com)  
Jennifer L O’Loughlin (jennifer.oloughlin@umontreal.ca)

**Version:** 2  **Date:** 14 October 2013

**Author's response to reviews:** see over
Reviewer: Jennifer Rose

Minor essential revisions:

- I would like to see the authors comment on extreme change in RMSEA for Time 2 for the test of factor mean invariance. Constraints typically worsen the fit of a model, which is apparent from the chi-square and CFI, but the RMSEA, and to some extent, the SRMR suggests a radical improvement in fit. This warrants some discussion.

The RMSEA value for the factor means at Time 2 is higher for the test of factor mean invariance, as would be expected. We corrected this typo in the Table and ensured that all reported results are accurate.

Reviewer: Shuqiao Yao

Major Compulsory Revisions

- At first, Data were drawn from an ongoing prospective cohort study of 1,293 adolescents, while the analytical sample included only 527 participants who provided complete data or had minimal missing data over follow-up. Although no statistically significant differences were observed for baseline depressive symptoms scores between adolescents included and not included in the analyses, participants included in the analyses were younger than participants excluded from the analyses, whether this age difference would have affected the results, the author should discuss it. Also, it is very important for authors to provide additional information, for example, whether there are differences on socioeconomic status, geographic locations (urban, suburban, and rural) and languages (French, English) between the original sample and the analytical sample.

We added (on page 7) that participants retained in the analyses were more likely than excluded participants to speak English (retained: 75.1% vs. not-retained: 57.1%, $\chi^2=42.40$, df=1, $p<0.05$). As a proxy for SES, we also added that the percentage of mothers (retained: 48.3% vs. not-retained: 41.9%, $\chi^2=3.77$, df=1, $p>0.05$) and fathers (retained: 48.5% vs. not-retained: 44.0%, $\chi^2=1.75$, df=1, $p>0.05$) who attended university did not differ significantly between the retained and not-retained groups. However, we cannot compare participants based on whether they lived in urban, suburban or rural areas as these data were not collected. Also, we added the following statement in the limitations section (page 15): “Participants retained for the analyses were marginally younger than those not retained and in addition were more likely to be English-speaking. It is possible albeit unlikely that these differences biased the results.”

- Some adolescents in this research speak English, while others speak French. I am quite confused that whether adolescents speaking English finished the English version of the scales, and adolescents speaking French finished the French version of the scales. Have the authors examined the equivalence of the two versions of the scales?

We added, on page 6, that French or English questionnaires were administered according to the language used in each study school. If questionnaire items were already available in French, we used these items. If not, the items were translated into French by two Francophone physicians, back-translated to test the accuracy of the translation, and then pilot tested in the target group for readability and ease of comprehension.
The factor structure of the DSS was tested at each time point using confirmatory factor analysis (CFA) with maximum likelihood estimation in LISREL. To use confirmatory factor analysis (CFA) with maximum likelihood estimation, the data should be normal distribution, and I am wondering whether the data in this study was really normal distribution. The authors should report this information. If the data were not normal distribution, maximum likelihood estimation should not be used.

Owing to the ordered-categorical nature of the data, the Satorra-Bentler scale chi-square statistic and robust standard errors were used since they yield unbiased goodness-of-fit indices when dealing with non-normal data. We clarified this in the manuscript (page 9).

Minor Essential Revisions:

- #CFI and #RMSEA values were not provided in Table 3 and Table 4, which makes it difficult for readers to know the exact differences between the nested models and the baseline model. Please add the #CFI and #RMSEA values in Table 3 and Table 4.

  The CFI and RMSEA value for each model are included in these tables. We added the change in CFI and RMSEA, as suggested.

- Please check the manuscript carefully. Some minor errors exist in this manuscript, for example, on page 4, line 9, “Hospital Anxiety and Depression rating scale [HADS; 20]” should be “Hospital Anxiety and Depression Rating Scale [HADS; 20]”.

  We made the corrections.

Reviewer: Marjolein Verhoeven

Major compulsory revision:

- The most serious issue I have concerns the way the authors examined time invariance. By using a multi group CFA framework, the authors treated the longitudinal data as if coming from three independent samples of adolescents. However, the authors are using longitudinal data and should treat it as such by incorporating the three waves into one statistical model to examine time invariance (see Motl et al. (2005) and Verhoeven et al. (2012))

  As suggested by the reviewer, we incorporated the three waves of data into one model to examine time invariance. We revised the methods section accordingly and now report these results (see Table 3). Also, we recognize that Motl and Verhoeven allowed residuals for all corresponding observed variables to correlate across time points in order to take the non-independence of observations into account when examining longitudinal invariance. However, modification indices in our study did not suggest that estimating these residuals in the models would improve model fit, meaning that the residuals did not correlate with each other to a statistically significant degree across time (likely because of the longer time period between assessments than in Motl and Verhoeven’s work). Thus, the residuals were not specified to correlate in the current study. We explain this in the revised manuscript. Of note, these re-analyses did not change the conclusion that the DSS is time invariant.
Discretionary revisions:

- In the introduction the authors argue that there is disagreement whether depression and anxiety are distinct constructs. It is therefore important to examine alternative factor models (resp. 1, 2, or 3 factors) for the DSS, as this questionnaire contains items regarding both constructs. It is, however, unclear why the somatic symptoms measured by the DSS should form a third factor. Some background information is needed here.

  *It can be difficult to determine if somatic complaints are expressions of a psychological distress and impairment as they may be attributed to biological and hormonal changes associated with puberty (e.g., hormonal fluctuations during menstrual cycles can cause fatigue, growth spurts can bring changes in appetite). Furthermore, somatic symptoms during adolescence have been shown to predict depression and other mental disorders during adulthood (Bohman et al., 2012). For both these reasons, it may be appropriate to distinguish between somatic and cognitive/affective symptoms. We added this justification in the revised manuscript (page 4).*

- Two recent publications regarding gender- and time invariance of depressive symptoms during adolescence should be included in the reviewed literature: Motl et al (2005) and Verhoeven et al (2012).

  *We now included both studies in the introduction and discussion of the revised manuscript (pages 4, 11, and 14).*

- In Table 3 it is shown that measurement invariance across sexes does not hold at all levels for Time 2 and Time 3. This is also the case for longitudinal invariance (Time 1 vs Time 2, and Time 1 vs Time3). Although the constraints of item uniqueness (model 4) and latent factor means (model 7) are indeed very strict tests of measurement invariance and are rarely met, I think the authors should still explain what the meaning of these results is for the DSS. Otherwise, why even investigate and report these strict tests?

  *In light of ongoing debate regarding the necessary level of invariance required to conduct meaningful comparisons, our stance is that a thorough examination of each level of measurement invariance is appropriate. For example, Meredith (1993) argued for strict invariance, while Vanderberg & Lance (2000) argued for strong invariance. Moreover, the meaning of each level of invariance is explained in the methods section (page 10 and 11).*

- It is not until the discussion of the manuscript that the authors let the reader know they used analytic techniques geared to ordinal values. This should be stated in the method section.

  *We have added a statement in the methods section (page 9) to inform readers that we used analytic techniques appropriate for ordinal variables. Specifically, we now state: “Owing to the ordered-categorical nature of the data, the Satorra-Bentler scale chi-square statistic and robust standard errors were used since they yield unbiased goodness-of-fit indices when dealing with non-normal data.”*