Reviewer's report

Title: The cross-cultural validity of the Resilience Scale for Adults: A comparison between Norway and Brazil

Version: 3  Date: 10 November 2014

Reviewer: Sigurd William Hystad

Reviewer's report:

This study explores the measurement invariance of the Resilience Scale for Adults across Norwegian and Brazilian samples. Resilience is an important topic and it is vital to have measures that can identify protective factors reliably and validly across cultures. Below are my comments and suggestions.

Major Compulsory Revisions:

1. Due to the number of t-tests conducted in comparing resilience scores across culture and sex, the authors use a p-value of .01 (line 340). What is the rationale for choosing this level (and not any other level)? More information is needed here.

2. Further, when presenting the result from the t-test, two p-values above the designated .01 level (p = .013, social resources in the Brazilian sample and p = .011, family cohesion in the Norwegian sample) are referred to as statistically significant. Are these p-values wrong or am I missing something here?

3. It is a bit unclear to me how the Satorra-Bentler Chi-square values were used in the model comparisons. Because the #S-B #2 is not #2-distributed it is not possible to use it in a standard difference test. Were the #S-B #2 corrected prior to using it to judge statistical significance? Please clarify.

4. The young age of the participants and the fact that they were all university students are discussed as a potential limitation to the generalizability of this study. I would add the unbalanced proportion of men and women to this equation. In both samples, but particularly in the Norwegian sample, there are considerably more women than men. How does this influence the generalizability of the findings?

Minor Essential Revisions:

5. The ms. should be thoroughly checked and scrubbed for any grammatical issues. Two examples:

Lines 387-388: As expected the RSA total score significantly negative correlation with HSCL-25 r = -.38, p < .01, and significant positive correlation with SOC r = .71, p < .01.

Lines 449-450: The RSA-factors were all significantly positive correlation with SOC.
6. The description of the test of differences in latent means on lines 180-184 is a bit unclear to me, especially the statement “If the kappa coefficients are non-significant following adjustment for variance in these properties, whereas a conventional regression analysis is significant, then sum score differences are…” Can the authors clarify what this means?

7. On lines 319-322 the authors state that the RMSEA and CFI index will be used to compare models in addition to the #S-B #2. I agree that this is a good idea, but it would be useful if these indices were actually referred to in the model comparisons. As it stands now they are simply included in Table 2 and not discussed at all. Also, some information on the substantive meaning of the differences in CFI and RMSEA would be useful (i.e., include some references with guidelines as to what is a “small” or “negligible” difference).

8. On lines 336-338 the authors report using Fisher transformations to assess the significance of the difference between correlations on the RSA-subscscales across the two samples. Why not include this element in the CFA invariance testing? Would it not make more sense to include an additional model that specifies equality constraints for the factor covariances?

9. On a similar note, I am not sure I see the necessity or the purpose of the t-tests testing differences in subscale means. Again, this element could more naturally be incorporated into the CFAs (and indeed already is). After all, the authors make the point that “the conventional practice of comparing two summatated raw scores using t-tests is based on the undue assumption that all of the above parameters (factor loadings, intercepts and residual) are equal.”

10. Explanations of the different parameter symbols should be included in the note to Table 2. It would make it easier to follow which and how many parameters are allowed to vary freely and which are constrained in the different models. Incidentally, the text on line 367 refers to model 4b, but there is no model 4b in Table 2. I guess this should be 4a?

Discretionary Revisions:

11. On lines 143-146, the authors present their aim to examine whether the RSA would exhibit “cross-cultural psychometric properties in a South American context.” I think that generalising a Brazilian (student) sample to a South American context is a bit of a stretch.

12. The discussion of the invariant intercepts on lines 417-425 is a bit underdeveloped. I realise that these are not the authors’ primary focus, but I would like to see some suggestions why they differ, besides the possibility that “the number on the scale mean different things for individuals in Norway and Brazil.” Are there any reasons to assume that the scales are interpreted differently across these cultures, and if so, what are these? Are there cultural differences between Norway and Brazil that leads to this particular pattern of differences (i.e., on average higher intercepts for the factors Planned Future and Social Competence, and on average lower intercepts for the factors Social
Resources and Family Cohesion in Brazil than in Norway)?

**Level of interest:** An article whose findings are important to those with closely related research interests

**Quality of written English:** Needs some language corrections before being published

**Statistical review:** No, the manuscript does not need to be seen by a statistician.

**Declaration of competing interests:**

I declare that I have no competing interests.