Author's response to reviews

Title: Factor structure and psychometric properties of the Trier Inventory for Chronic Stress (TICS) in a representative cross-sectional German survey

Authors:

Katja Petrowski (katja.petrowski@tu-dresden.de)
Elmar Brähler (elmar.braehler@medizin.uni-leipzig.de)
Sören Paul (soeren.paul@uniklinikum-dresden.de)
Cornelia Albani (cornelia.albani@schussental-klinik.de)

Version: 2 Date: 8 November 2011

Author's response to reviews: see over
Dear Mr. Aldcroft,

Thank you very much for considering our manuscript “Factor structure and psychometric properties of the Trier Inventory for Chronic Stress (TICS) in a representative cross-sectional German survey” by Katja Petrowski ¹, Elmar Brähler ², Sören Paul ¹, & Cornelia Albani ². We also thank the two reviewers for their comments. In the following, we have addressed the comments. Please note that the reviewers’ comments are typed in italics.

We look forward to hearing from you.

Sincerely yours,

Katja Petrowski, PhD

Authors’ affiliation:
¹ corresponding author (Tel.: 0049 351 - 458 3634, Fax: 0049 351 - 458 6318, Mail: katja.petrowski@tu-dresden.de )
² Dresden University of Technology, Department of Psychotherapy and Psychosomatic Medicine, Fetscherstrasse 74, 01307 Dresden, Germany
³ University of Leipzig, Department of Medical Psychology and Medical Sociology, Philipp-Rosenthal-Straße 55, 04103 Leipzig, Germany
Editorial Requests:

1. We recommend that you ask a native English speaking colleague to help you copyedit the paper. If this is not possible, you may need to use a professional language editing service. For authors who wish to have the language in their manuscript edited by a native-English speaker with scientific expertise, BioMed Central recommends Edanz (www.edanzediting.com/bmc). BioMed Central has negotiated a 10% discount to the fee charged to BioMed Central authors by Edanz. Use of an editing service is neither a requirement nor a guarantee of acceptance for publication. For more information, see our FAQ on language editing services at http://www.biomedcentral.com/info/authors/authorfaqs#12.

   We sincerely regret that our manuscript did not fully satisfy the language requirements. We have therefore revised it entirely and then, following your suggestion, presented it to a native speaker for copy-editing. We are hopeful that the manuscript has now been improved to the degree of making it acceptable for publication.

2. In the background section of your abstract, please include a statement of your research aim.

   We now explicitly name the aims and hypotheses of our study.
Reviewer #1

Major Compulsory

1. Let me preface my comments by suggesting the authors have a native English speaker review the manuscript. I worry that some of the concerns I will raise can be attributed to language issues. I will also work through the document page by page although I will try to highlight the more important concerns.

We are sorry that the presentation of our study did not fully satisfy language requirements. Therefore, we have reworked the whole manuscript and tried to make the statements easier to follow by omitting or replacing misleading statements based on the reviewers’ suggestions. Finally, a native speaker was consulted for proofreading the text. Hopefully, it now meets the requirements for publication.

2. Page 5, first paragraph: Because the TICS is based on a theoretical model, “the authors automatically saw content validity as a given”? That's a very odd way of putting it. And then the next sentence says that “Factor-analytical analyses showed good factorial validity”. That's the point of this article as I understand it so far more needs to be said about that than this one statement. Especially given the next sentence says the factorial structure has not been replicated. So then how do the authors conclude the measure has good factorial validity?

We understand that it was hard to follow these short statements. Therefore, factorial validity, configural invariance, and measurement invariance have now been discussed in more detail (see page 4).

Furthermore, if factorial validity has been established, then I'm not clear as to why an EFA was required at all. The authors should have been able to move directly to a CFA.

This is correct, and the sentence may have been misleading. Factor validity still needs to be established on a broader basis, i.e. with representative samples. Therefore we chose to take a more conservative approach first looking for possible factor structures that might emerge despite the nine-factor structure that had been published by the EFA. Then, the main part, the confirmatory factor analyses were carried out with the emerged models, including tests for measurement invariance. We thought this two-step procedure to be more persuasive than simply comparing model fits, especially in case t we would find the hypothesized nine-factor model – which we actually did. If you still recommend it, however, we could omit the EFA and just report the CFA results.

3. Page 9, second paragraph: I’m lost in the EFA. The abstract speaks to a four factor structure. I don’t see that in the results section! Rather, a solution between “seven to nine” is presented when there are eight eigenvalues greater than 1? However, the authors then move to a two factor model, albeit a model with 9 cross-loading items? I don’t understand the jump from 7-9 to 2 factors.

Furthermore, the amount of variance accounted for by the two factors is small.

We are sorry for the brief EFA results section. The presentation of these results has now been extended. We hope this makes the section more explicit and easier to follow (see page 8-9).

The four-factor solution was an unfortunate mistake which has now been removed.

The reason for choosing 7, 8, or 9 factors has now been explained in more detail (see page 8). Unfortunately, to our knowledge there is no rule on a confidence interval that is acceptable around the 1.00 cut-off, so we had to choose the number on theoretical assumptions, thus preferring the published nine-factor solution.

In fact, we tested all the emerging meaningful solutions in the CFAs. We tested (a) a two-factor solution as suggested by the Scree test and stated by the TICS authors, (b) a nine-factor solution as suggested by the Kaiser-Guttman-criterion and the TICS authors, and (c) a combined nine-plus-two factor model. A null-model with only one main factor has been omitted from the study.

On the one hand, the number of complex items illustrates that two-factor solutions with smaller/larger δ-values result in fewer simple structures with more hyperplane or complex items. On the other hand, these numbers and the small amount of explained variance demonstrate that a solution with two factors might not be the best one possible. This made us realize that
that the nine-factor solution was the superior one.

4. Page 10: I’m not clear why a one factor solution is being tested. Yes, the nine-factor solution is a better fit than the two factor solution but the difference isn’t very great--- isn’t there something to be said for parsimony? Thank you for these queries. We no longer report this model, however, we would do so in case you find it necessary. As for the comparison of the two-factor and the nine-factor model, we believe the differences in fit indices to be important. In the statistics presented in the first submission, the nine-factor model had only 2/3 of the Chi² of the two-factor model, CFI and TLI are .10 greater, respectively, the RMSEA is not acceptable but good, the SRMR is 9% lower, and the AIC is also 2.5% lower. Therefore, every index in the nine-factor model was more favorable, including indices that are sensitive to parsimony (e.g. RMSEA). New CFA results have now been included, comparing (a) a model with two factors to a model with nine factors, (b) a model with nine factors to a model with nine-plus-two superior factors, and (c) models with more or less restrictive constraints on parameter estimations, i.e. tests of measurement invariance.

More importantly, I’m not sure why the authors conclude there is moderate to good fit for any of the models. TLI and CFI for all three solutions are very poor and far below what the authors specified as acceptable earlier in the manuscript.

Thank you for detecting this incorrect representation. The indices do indeed reach from unacceptable (all CFI, TLI) to acceptable (RMSEA) to good (SRMR). We have corrected this. Furthermore, according to Marsh, Wau, & Zen (2004), the recommendations by Hu & Bentler (1999) are hardly met by questionnaire data, especially when they have many items. In our case, the questionnaire consists of 57 items on nine scales. Furthermore, Marsh and colleagues argue that goodness-of-fit indices may be more appropriate for comparing models with each other instead of judging the model fit itself. This limitation has now been added to the text (cf. page 11).

5. The authors need to provide somewhere the items, their loadings on ALL of the factors (with the loadings for the assigned factor) in bold. Otherwise, I do not believe the reader can verify (or make sense) of the results here. That would be far more useful than what is presented in Table 1 or the figures.

This is a very good annotation and makes an excellent addition to the data already presented. Since publishing all the items in English might produce copyright problems, we have added item examples for each factor (see pages 5-6). The recommended table has now been included in the manuscript and, hopefully, will help the reader to make sense of our findings.

6. Page 11, I’m not sure what the authors mean by a two-way covariance analysis for age and gender? I’m assuming the authors compared genders and ages and education on the subscales of the TICS. But those would be comparisons of MEANS. The authors need to demonstrate FIRST that the measure is invariant across genders and ages and education --- that the factor structure is the same. Only then is it appropriate to compare means --- that should be the last thing tested, not the first and not the only thing. The authors have the sample size to do this, especially if they do not do the EFA or if they combine the EFA and CFA data to do tests of invariance.

The ANOVA was calculated with age group (cf. Table 3) and gender as dependent variables, skipping education as an additional dependent variable, and the nine TICS scales as independent variables, each. These results are presented and interpreted (see page 10). Also, we have adopted your excellent suggestion of doing tests of measurement invariance via CFAs.

Minor Essential
1. Page 3, Introduction, first paragraph: I’m not clear what is meant by “as within two years” in the description of stress research. Two years from what? The appearance of a stressor?

This section has been rephrased (see page 3).

2. Page 3, Introduction, second paragraph: I found the differentiation of chronic and acute stressor to be confusing. A definition is provided for chronic stress; no
such definition is provided for acute stress. Furthermore, the definition of chronic stress is unnecessarily complicated. The first part of the definition speaks to repeated stressors or the absence of mechanisms to deal with stressors, but then we’re told chronic stress may also start slowly (compared to acute stressors?), last for a short or long time (but if it is a short time, is that chronic?) and it cannot be controlled (well yes, but that’s true of acute stressors too, and it is not so much whether they can be controlled but whether the response to the stressors can be moderated). And then satisfaction with one’s needs is added to the mix as well as acute stressors transforming into chronic stressors. I’m left thinking pretty much everything is a chronic stressor. And then the next paragraph introduces “experienced stress”; I’m not sure what that is.

This very good remark has helped us to describe the construct more stringently. Hopefully, the definition has now been clarified and any confusing wording been eliminated (see page 2).

3. Page 3, Introduction, fourth paragraph: We’re presented with the TICS. It is described as an area-specific instrument. I’m not sure what area-specific is in reference to? Later it is said that the “TICS focuses much more on stressful life experiences followed by chronic stress…and on area-specific stress reactions” compared to “other stress questionnaires” but I’m not familiar with the measure and the items for the measure are not provided so I’m not sure what that means.

We are sorry for having caused confusion; we have now clarified the statements (see page 3).

4. Nine dimensions (subscales?) are listed but it is not clear to me why those nine dimensions. I take it those dimensions are drawn from the systematic requirement-resource model of health, but I’m not familiar with that model. A very brief description of this model follows and we’re told that model is captured by two TICS scales --- lack of need satisfaction and stress due to high demands --- but I’m not clear how those two “dimensions” related to the “nine dimensions” of the TICS.

This well-taken remark points to the lack of clarity in our manuscript. The core points of the manuscript are: the nine dimensions that were originally found with item-based EFA we tried to replicate. The two factors you cited were found by the authors of the TICS with sum-score-based EFA. We tried to show that a model combining both the structure of the items (nine dimensions) and the structure of the factors (two higher-order factors) results in an even better fit within a representative sample. However, the addition of the two higher-order factors did not substantially change the fit indices. Thus, the advanced model is reported as a heuristic model with practical relevance (see pages 12-13).

5. Page 7, second paragraph: I’m confused by (a) the statement “at least one item of each scale must be completed for the questionnaire to be examined”, (b) why missing items were replaced by scale means --- that’s not a good way to handle missing data, (c) but then the statement “only completed questionnaires were included in the factor analyses” --- so there were no missing data?

This misleading “statistical procedure” section has been edited extensively. Now, we no longer use the missing data strategy (a) in which at least one item per scale had to be answered so that the other items on the scale could be estimated. Else, the entire questionnaire had to be abandoned. (b) We have now switched to using listwise (scalewise) exclusion resp. the calculations, i.e. (c) if any one of the 57 items of the questionnaire is left unanswered, the entire questionnaire is not used in the factor analysis.

I do commend the authors for separating their large dataset into two samples --- not sure why they are called “partial samples”, though. Later we are told missing data are excluded “list by list” --- I don’t know what that means? Listwise?

Thank you very much for the acknowledgment. We called it partial samples in order to make it easier for the reader to follow with the differing N numbers, since this procedure seems to be unfamiliar to some readers. Yes, this meant listwise. This sentence, however, has been omitted since it is clear that only complete questionnaires were used in ALL the factor analyses (EFA, CFA) – that is, the listwise exclusion mentioned earlier in the “sample” section.
6. I’m not familiar with Oblimin rotation but I would have thought one would have gone with a fixed value rather than varying the parameter, but again I’m not familiar with that rotation. And why that rotation and not some other oblique rotation? A brief explanation here or elsewhere would be good.

Looking for the simplest structure by examining solutions with different delta values is recommended in several guidelines, e.g. Gorsuch (1983, pages 189-190 and 202-206) or at [http://faculty.chass.ncsu.edu/garson/PA765/factor.htm]. Oblimin is the oblique rotation chosen in order to use the same methods as the TICS authors and is one of the standard rotations according to Gorsuch (1983). This information has been added to “statistical procedure” (see page 7).

7. Page 9, first paragraph: “Items were appraised by 55% to 84% of the participants”. Does appraised mean completed? And why such low numbers? My guess is that these items are specific to events that might occur in one’s life, but again without having access to the items, I’m not sure.

Thank you very much for this insightful remark. The misleading sentence has been corrected. The numbers stated represent the number of yes-answers (in numbers: 1, 2, 3, or 4), which has now been pointed out in the text (see page 8).

8. When conducting the CFA, the authors allow no cross-loading items --- but are there not nine complex items according to the EFA?

That is correct, and it may have affected the fit indices of this solution in the CFA. We now report revised results.

Discretionary

1. Page 4, second paragraph: Reliability is NOT a property of a scale but rather a characteristic of a scale in regards to a particular sample.

Very true, therefore we have now made mention of it (see page 6).

2. Page 6, first paragraph: I do not know what a random-route sampling method is?

Concerning this question we have included an explanation in the manuscript (see page 5). The random-route sampling method is characterized as “national samples are based on multilevel selection processes. Multilevel selection processes are a combination of random or/and systematic and/or stratified probability samples at different levels: stratified selection of sampling units, systematic selection of households by random walk, random selection of one person per household by Kish-table.” For further details you might consider: Hoffmeyer-Zlotnik JHP: New Sampling Designs and the Quality of Data. Developments in Applied Statistics 2003. [http://mrvar.fdv.uni-lj.si/pub/mz/mz19/hoff.pdf]

3. Page 6, last paragraph: If I am correct, the chronic stress screening scale takes 12 items from some of the other scales. How does that work? I do not know what to make of that screening scale and what role it should or does play in subsequent factor analyses if any? Can it be omitted for the purposes of establishing factorial validity of the TICS?

Our intention was not to validate the screening scale by using EFAs/CFAs but rather to show that the scale is related to most of the now validated scales. It has now been omitted from the manuscript.

4. Minor thing --- p is never really .000, there is always some chance of a Type I error so should be p = .001.

Of course, we have corrected all the p-Values accordingly.

5. Page 10, last paragraph: Why split-half reliabilities as opposed to Cronbach’s alpha?

Both reliabilities are calculated differently and represent different estimators of the true reliability. ALPHA can be understood as the lower bound for the true reliability, especially when the items produce differing factor loadings (McDonald RP: Test theory: A unified treatment. Mahwah, NJ: Lawrence Erlbaum Associates; 1999:p.92). Split-half reliabilities are therefore closer to the real reliabilities.
Reviewer #2

This a a psychometric paper on TICS. In general the paper merits publication but before any decision the following changes should be made:

Abstract:
1. Remove reference from line 7.
   We have removed the reference.
2. Avoid using the word subject.
   The wording has been changed.
3. Remove the last sentence from conclusion.
   This sentence has been omitted.

Background:
It seems that the paragraph before the last paragraph (from line 11 page 4 to the end of this paragraph) should be removed from here.
   (a) Scale reliabilities are now stated in the “Instrument” section (see page 6) to be compared to the reliabilities we found in our sample. Types of validity other than factorial validity (introduction section, fifth paragraph) have now been omitted due to your helpful comment.

Results:
1. EFA: If the criteria is to accept eigenvalues of 1 or greater, then it is not clear why the authors report on values of 0.93 and 0.87. In fact, it seems a 8-factor solution emerged. This should be clarified. In addition where is a four-factor solution (you mentioned in the abstract but it does not appear hear)
   The somewhat brief presentation of the EFA results has now been extended and rewritten to make it easier to follow (see page 8). The reason for choosing 7, 8, or 9 factors has now been stated. Unfortunately, to our knowledge there is no rule on a confidence interval that is acceptable around the 1.00 cut-off so we had to choose the number on theoretical assumptions. We have presented all the values above 1.00, the subsequent one that is still around 1.00, and the one value following it, which is just to illustrate the further decline of eigenvalues. The four-factor solution is a regrettable mistake that has now been remedied.
2. The examination of CFA needs another sample. This should be discussed as limitation. You need report on different models based on EFA. Please clarify what was your intention using present models.
   You are absolutely right. Hence, we argued in the same direction (see page 7). Our solution was the following: the sample of $N = 2,339$ participants had been split by random into equally distributed partial samples for the EFA ($n_{EFA} = 1,190$, $n_{CFA} = 1,149$). The two partial samples did not differ in their distribution of important variables as tested with ANOVAs (see page 7, statistical procedure). The limitation of this procedure has now been addressed in the new limitations section at the end of the discussion (see below). Furthermore, reviewer #1 has made a commendable remark on this procedure. Thus, for sample-splitting, its presentation and discussion we hope to have complied with your requests for clarification. If still recommended, the EFA can be omitted and only the CFA results reported. Calculations based on $N=2,339$ would, however, not result in substantial changes in the fit indices!

Discussion:
1. You need to add a limitation section and report on limitations.
   Such a section has now been added (see pages 14-15).
2. You need to include a conclusion section.
   Thank you for pointing this out. The very last discussion paragraph has now been adapted to precisely state the conclusions drawn. (see page 15).
This mistake has been deleted.