Author's response to reviews

Title: Positive versus negative mental health in emerging adulthood: A national cross-sectional survey

Authors:

Regina Winzer (regina.winzer@folkhalsomyndigheten.se)
Frank Lindblad (frank.lindblad@neuro.uu.se)
Kimmo Sorjonen (kimmo.sorjonen@ki.se)
Lene Lindberg (lene.lindberg@ki.se)

Version: 2 Date: 14 November 2014

Author's response to reviews: see over
Dear Dr Victorino Silvestre

We thank you very much for commenting on our manuscript. Further, we appreciate the reviewer’s proposals to improve the quality of the article. We have addressed them all, made the recommended changes and additionally had the manuscript copy-edited by Edanz. As these corrections have been substantial and our manuscript has undergone new statistical analyses we send you a new, clean copy. We now hope the article has reached the standard to be published in your journal.

In our point-by-point-response below we quote each reviewer’s comments (italicized), reply to them sequentially and refer to the changes we made.

Reviewer Helen Winefield:

Minor Essential Revisions
There are a few mistakes in the English expression, which although minor, damage the readability:
1. p. 9, plural is indices, not indexes
   We have amended the plural form to “indices” in two cases, p. 9 last paragraph, second last line and p. 10, line 2.
2. p. 10, "contrarily" should be "in contrast"
   The mistake is corrected, Result section, p. 10, second paragraph, second last line.
3. p.11, "extincted" is not a word. Perhaps use "extinguished". Or say the differences disappeared.
   We corrected the error in two cases, Result section, p.12, line 6 (“extinguished”) and ibid line 8 (“disappeared”).
2. p. 11, 2nd to last line: say "being aged between 16-29 years".
   Changed to (Discussion section, p.15, line 1): “Our data indicate that PMH decreases as age increases in people between 16 – 29 years old, …”
3. P.15, line 5 of 2nd paragraph: "opportunity to assess”.
   We omitted this sentence as we considered it would not contribute to the manuscript.

Discretionary Revisions
It seems to me that the conclusion (p.16) that it is still an open question whether GHQ captures two dimensions of mental health, is rather weak given the data. The clear clustering of positively vs negatively worded items suggests to me that the factor analysis results do not convincingly reflect bi-dimensionality, but are indeed an artefact of the wording. That is supported by the analyses of associations between PMH and NMH and the potential predictors of health. A stronger concluding statement about the apparent unidimensionality of the mental health concept would therefore be appropriate.

We have according to reviewer 2 made the following changes:
- Abstract, p. 3, line 1 – 3: “Our results could be seen to indicate that positive and negative mental health are distinct and complementary constructs. Rather, the results of our factor analysis may specifically reflect the wording of the items.”

- In the manuscript, Conclusion section, we eliminated the first sentence “It is still an open question whether GHQ-12 captures two dimensions of mental health/ill health or not.” The new formulation is as follows: “Ultimately, we were unable to determine whether GHQ reflects two mental health dimensions. We cannot rule out the alternative hypothesis that the fit of the two-factor model is related to the wording of items. Based on these findings, we suggest that the GHQ-12 is useful for the purpose for which it was intended, i.e., to detect “psychiatric morbidity” in clinical and population settings. We believe that future investigations of the dimensionality of “mental health/ill-health” needs to use instruments that are specifically adapted for this purpose.” (Conclusions, p. 15 last paragraph last 2 lines - page 16 line 1 – 5.)

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

In accordance with the editor’s advice the manuscript has now been copy-edited by Edanz, a provider recommended by BMC.

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

According to the advice of reviewer 2 a statistician has scrutinized the manuscript and re-analysed the data by a latent trait approach conducted with Mplus.

**Reviewer Tim J Croudace:** 1

**Major Compulsory Revisions:**

1. In the abstract one aim/objective described is “to explore if dimensions corresponding to positive and negative mental health could be identified in a sample of young individuals”. These dimensions are not established in this paper. A traditional PCA is applied (as if the data were continuous, whereas they are actually ordinal) and then one of two derived scores is called “positive mental health”. This alone does not establish anything positive re: mental health. CFA/SEM is also applied, without using an approach suitable for ordinal items. Although lots of the references cited adopt a strong methodology for analysis of ordinal GHQ item responses, the paper does not: this is a major weakness that could be addressed. The dominant methodological advice cited by the authors and available in the literature for over 20 years, is to use a latent trait approach to the GHQ data, when reporting on dimensions or factors. There is also advice on scoring methods that might be relevant too. Latent trait analysis, or categorical data factor analysis (multidimensional item response modelling) is well documented, available in free software (e.g. FACTOR) and should be used in order for any of the claims re dimensions or predictors of dimensions to be made.

Our data have now undergone additional new analysis conducted with Mplus 7.11 software. The PCA has been substituted by an EFA to fit ordinal data. Our new formulation is as follows, p. 9 last paragraph and p.10 line 1 – 6: “Data analyses were conducted with Mplus 7.11 software. To assess whether the GHQ-12 could be used to measure both positive and negative mental health, we used exploratory factor analysis (EFA). Scree plots, eigenvalues,
and differences in model fit were used to determine the number of factors. We performed orthogonal (varimax) and oblique (geomin) rotation analyses. We also conducted confirmatory factor analysis (CFA). We used two indexes of incremental fit—the Tucker–Lewis index (TLI) and the comparative fit index (CFI)—to assess the fit between specified models and the data. We considered a value > .95 on these indexes to indicate a close fit. We also used the root-mean-square error of approximation (RMSEA) on which a value < 0.05 is considered to indicate a close fit [40]. The GHQ-12 items were defined as ordinal and we used a robust weighted least squares estimator. We used the DIFFTEST function in Mplus to calculate the significance of the difference between the effects on positive mental health and negative mental health.

The data analysis resulted in a new Table 2, which replaces the former Table 2, p. 26.

We now conclude from our analysis: “We used a CFA to verify the factor structure. A model where all items were treated as indicators of the same latent variable was a significantly worse fit to the data compared with a model that had two separate latent variables, \( \chi^2(54) = 43,183, p < .001; \) TLI = .894; CFI = .913; RMSEA = .139; and \( \chi^2(53) = 12,492, p < .001; \) TLI = .969; CFI = .975; RMSEA = .075, respectively. The fit of the model with two separate latent variables (Figure 1) could be characterized as quite satisfactory according to the criteria presented by Hu and Bentler (1999). The factors were tentatively called “positive mental health” (PMH) and “negative mental health” (NMH).

” (P.10 last paragraph – p. 11 line 1 - 3.)

2. It is not the case that showing differential predictors provides evidence for different dimensions. This aspect of your methodology could be discussed with a statistician and some rationale for this being a valid method for establishing that goal of analysis revisited.

We have discussed this issue with a statistician who helped us to conduct the data analyses shown above. If we understand this comment correctly we have done the following, presented new results in Table 3, p. 27 which replaces the former Table 3. Based on them we now conclude: (Discussion section, p. 13, 2nd paragraph line 1-8)

“The results of our factor analysis thus seem to support the hypothesis that positive and negative mental health are different constructs. This is consistent with earlier findings [4, 20]. On the other hand, the mirror-like associations between these constructs and the potential predictors fit well with a unidimensional model. It is possible that the results of the factor analysis specifically reflect the wording of the items, rather than influence of the two hypothesized health constructs. Several authors have suggested this interpretation. For instance, Hankins claimed that the bi-and multi-dimensionality attributed to the GHQ [41, 42] is an artefact because of a response bias of the negatively worded items…..”

Minor essential revisions:

1. Please retain the ordering of GHQ items in your tables: number them using item numbers.

In Table 2 the GHQ items now have been numbered by their item numbers, p. 26.
2. Response options for the GHQ are verbally anchored statements (not numbers). Only at the point of scoring is an ascending integer series (1,2,3,4 or 0,1,2,3) applied in Likert style. Hence the response scale is not Likert really since the response options are worded differently across positively and negatively phrased items. More accuracy is required please in the description. There are many poor and inaccurate descriptions published: it would be good and easy to improve on this situation by reporting exactly what GHQ responses are. There are no numbers on the GHQ form. The numbers are values applied at the time of data entry and often then used eventually as variables for analysis (ordinal item response scores). Summing the 1234 values is called Likert scoring, but models for 1234 responses are usually given a modern treatment as ordinal variables, even in PCA or factor analysis (CFA).

We have specified our section Study variables, p.7, 2nd paragraph in this way: “The items included in the GHQ-12 are listed in Table 2. Items 1, 2, 4, 5, 7, and 12 had four options: ‘More/better than usual’, ‘Same as usual’, ‘Less/worse than usual’ and ‘Much less/much worse than usual’. Items 3, 6, 8, 9, 10, and 11 had the following four options: ‘Not at all’, ‘Not more than usual’, ‘More than usual’ and ‘Much more than usual’. The responses were coded as ordinal variables”

3. If the authors want to retain a GHQ 1234 responses as continuous variables approach, then they should provide the rationale and justification, and probably a sensitivity analysis. This would probably require some CFA comparison based on estimators such as MLR vs WLSMV and use of something like, or equivalent to the functionality available in Mplus.

The GHQ responses have been amended and are now treated as ordinal and a robust weighted least squares estimator (WLSMV) was used.

**Discretionary Revisions:**

1. You should mention that the GHQ was developed to screen for minor psychiatric morbidity. If it is able to measure anything that might be termed “positive” in relation to mental health then that is due to the design decision (at the point of building the GHQ questionnaires) that incorporated some guidance on the possible value or need for balanced questions i.e. some positively phrased questions and some negatively phrased. For the GHQ this means that some of the versions (12, 30) have half of their items phrased either positively or negatively but the response options (the verbal anchors on the copyrighted forms) are also reversed. This is importance since balanced questionnaires do not necessarily reverse their response options. There is therefore no need to reverse score the GHQ items. Someone who is very distressed and (in theory could) indicating the most distressed level on each item would therefore still complete the rightmost response option on the copyrighted form. Hence the balanced design does

In our Abstract, p. 3, line 3 we made the following amendment: “We conclude that the GHQ - 12 is an appropriate tool for its original purpose, to detect “psychiatric morbidity”. The respective amendment has been made in the manuscript, Conclusion section, p. 16, line 1:
“Based on these findings, we suggest that the GHQ-12 is useful for the purpose for which it was intended, i.e., to detect “psychiatric morbidity” in clinical and population settings.”

We appreciate your comments and thank you once again.

Best regards,
Regina Winzer