Reviewer's report

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

Version: 1  Date: 30 August 2014

Reviewer: Tim J Croudace

Reviewer's report:

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.

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.

Minor Essential Revisions:

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

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).

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.

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 not rule out response effects due to the layout of the response options.

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

Quality of written English: Acceptable

Statistical review: Yes, and I have assessed the statistics in my report.

Declaration of competing interests:

I have calibrated an educational test for GL Assessment who distribute the GHQ.