Jornal de

Pediatria

LETTER TO THE EDITOR

An illustration of how harmful the degree of freedom of the researcher is using the article entitled ''Reliability and validity of the Brazilian version of the Pittsburgh Sleep Quality Index in adolescents'' as an example

Dear Editor,

I recently had the opportunity to read the article entitled ''Reliability and validity of the Brazilian version of the Pittsburgh Sleep Quality Index in adolescents'',[1] published in 2016 after a proper peer-review process. This article had come to me because of my work on a project in which the Brazilian version of the PSQI was utilized to obtain information related to sleep in psychiatric patients, granted by a public initiative.

However, after reading the article, its analyses, and the conclusions derived from it, I could identify several issues that I'll briefly describe in the following paragraphs. I'll use the bullet points to address each point in a clear way.

1 The authors point out that "The validity of the PSQI components was done through an exploratory factor analysis, with orthogonal varimax rotation, with a sample of 209 adolescents."

Unfortunately, the authors do not make the difference between factor analysis and principal component analysis clear. Both methods are different, which is well-documented in international literature,[2,3] but also by Brazilian researchers.[4–7] Considering the text written, it seemed to me that the authors chose a PCA solution as they carried out a ''varimax'' rotation.

Unfortunately, the choice of this rotation was deeply biased and I will list two reasons to justify my sentence. First, orthogonal rotations means that the angles of all factorial axes were held constantly at 90° during rotation. Therefore, the factors are forced to be unrelated. There are multiple pieces of evidence pointing out that psychological factors are correlated, which justify the use of oblique solutions.[8] However, in the second part of their study, the authors just changed this orthogonal solution with an oblique solution, also confirming my sentence.

2 ''The value of the Kaiser - Meyer-Olkin sample adequacy measure was 0.59,''. This result is of interest. Looking at this test through a statistical lens, the KMO compares the magnitudes of the observed correlation coefficient (Product-moment) in relation to the magnitude of the partial correlation coefficient. ''Low'' values for KMO suggest that the data set isn't well suited for factoring, and that 0.7 is the most-often recommended criterion.

That said, it's crystal clear that the solution derived was inadequate in the preliminary tests.

3 The Bartlett-chi-square sphericity test had an approximate $2nd = 382,992$ ($p = 0.000$). The authors report that $p = 0$, but this does not exist. Statistics is about probability, therefore the correct value is $p < 0.0001$.

4 Table 1 of the article indicates that the factors were not well identified. A factor can be interpreted if one can find ''at least 3 non-cross-loading items with an acceptable loading score''. In addition, one should keep in mind that ''Following the advice of Field (2013: 692) we recommend suppressing factor loadings less than 0.3''.

With that said, the factor was formed of 2 items in which the factorial loads are greater than 0.3. Factor 3 presents items with cross-loads greater than 0.3. One will not be surprised with this result, as this failure was previously granted by the KMO results.

5 According to the authors: The PSQI obtained a high internal consistency, with Cronbach's 0.71. However, I should bring to the debate that Cronbach's Alpha assumes tau-equivalence.[9] When one only computes the ''Total'' Cronbach's alpha, the result obtained is fundamentally flawed. Moving forward with my sentence, it is well documented that Cronbach's alpha needs to be > 0.80 in exploratory research, and that each construct needs to be assessed, with results above 0.7 regarded as just marginal.[10]

6 There was a statistical difference between the PSQI scores in the test and retest ($p < 0.001$) which indicated the presence of a systematic error, confirmed by the analysis of the Bland - Altman plot.

The authors were correct here. This outcome indicates that the results are not reliable. Unfortunately, the title of their manuscript, the abstract, and the conclusion makes the opposite claim.

7 The confirmatory factor analysis image makes it clear that modification index (MI) techniques were used to ''try'' to fit the model. This is clear from the covariance observed in items 6 and 5. I stress that this was not documented in the ''analysis section''. In addition, MI is highly problematic and criticized among psychometricians and statisticians. Some suggest that this method is either a form of data hacking or a devious departure from the theory-driven paradigm.[11]

8 The conclusion ''Finally, the Brazilian version of the PSQI demonstrated high internal consistency and moderate reliability in adolescents. The original version of the instrument proved to be valid for evaluating sleep disorders in adolescents, between the model composed of two factors, excluding the component on the use of sleeping medications, obtained better adjustment values, it seems to be the most adequate to assess the different characteristics of sleep in this population'', is not true. The authors tortured the statistics until they were able to demonstrate the adequacy of the instrument, which unfortunately went unnoticed by the magazine and reviewers. The famous, anonymous, and old quote is well-posed there as an empirical demonstration ''If you torture data long enough, it will confess.''

Before my conclusion, it is worth saying that ''Validity refers to the degree to which evidence and theory support the interpretations of test scores for proposed uses of tests''.[12] Therefore, validity is not a test property!

Science is remarkably self-correcting and I appreciate the opportunity of post-reviewing this manuscript. There are no conflicts of interest. The good faith in the improvement of research in statistics and psychometrics motivated me in setting out this letter.

Thank you!

Conflicts of interest

The author declares no conflicts of interest.

References

1. Passos MH, Silva HA, Pitangui AC, Oliveira VM, Lima AS, Araújo RC. Reliability and validity of the Brazilian version of the Pittsburgh Sleep Quality Index in adolescents. J Pediatr (Rio J). 2017;93:200–6.
2. Kim HJ. Common factor analysis versus principal component analysis: choice for symptom cluster research. Asian Nurs Res (Korean Soc Nurs Sci). 2008;2:17–24.
3. Widaman KF. Common factor analysis versus principal component analysis: differential bias in representing model parameters? Multivariate Behav Res. 1993;28:263–311.
4. Anunciação L. An overview of the history and methodological aspects of psychometrics-history and methodological aspects of psychometrics. J Reatt Ther Dev Divers. 2018;1:44–58.
5. Damásio BF. Uso da análise fatorial exploratória em psicologia Usos de análisis factorial exploratoria en psicología. Avaliação Psicológica. 2012;11:213–28.
6. Primi R, Muniz M, Nunes CHSS. Definições Contemporâneas de Validade de Testes Psicológicos. In: Hutz Cláudio Simon, editor. Avanços e polêmicas em avaliação psicológica. São Paulo: Casa do Psicólogo; 2009. p. 243–65.
7. Golino HF, Epskamp S. Exploratory graph analysis: a new approach for estimating the number of dimensions in psychological research. PLoS One. 2017;12:e0174035.
8. Costello A, Osborne J. Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. Pract Assessment Res Eval. 2005;10:1–9.
9. Raykov T. Scale reliability, Cronbach's Coefficient Alpha, and violations of essential tau-equivalence with fixed congeneric components. Multivariate Behav Res. 1997;32: 329–53.
10. Yurdugül H. Minimum sample size Cronbach's coefficient alpha: a Monte-Carlo study. Hacettepe Egit Derg. 2008;35:397–405.
11. Steiger JH. Structural model evaluation and modification: an interval estimation approach. Multivariate Behav Res. 1990;25:173–80.
12. AERA, APA, NCME. Standards for educational and psychological testing. American Educational Research Association; 2014.

Madalena S. Santos [iD]

Universidade Federal do Paraná, Curso de Estatística, Curitiba, Curitiba, PR, Brazil
E-mail: madalena.gestora@gmail.com