



## Answer to the Letter “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,

We carefully evaluate the observations made by the author of the Letter to the Editor.<sup>1</sup> We appreciate the opportunity to respond because we believe that this type of debate contributes to the maturation and advancement of knowledge. Below are the answers to the main points highlighted in your letter.

In our study, we performed the extraction method using the principal component analysis, which is one of the methods that is commonly used to produce linear combinations to allow the explanation of the maximum variability of the data. We opted for this method based on the recommendations of Tabachnick and Fidell,<sup>2</sup> who describe that this method of extraction should be used when the interest of the research is to summarize variables in smaller sets and to help in the definition of the number of factors. Regarding the choice of the rotation method, we understand that it is a controversial topic in the literature used. Some authors<sup>3,4</sup> argue that the orthogonal and oblique rotation methods produce similar factorial loads. Fávero et al.<sup>4</sup> argues that from a theoretical point of view there is no analytical reason to favor one method over another, and highlights that rotation does not affect the quality of the suitability of factor model, the commonality, and the total variance explained by the factors. It really happened. There was no significant difference in the results when we tested the two rotation methods. Thus, we chose to use the varimax method that maximizes the variation between the weights of each component and makes the groups of factors more interpretable.<sup>4,5</sup>

Regarding the adequacy of the sample, we initially considered the recommendations in the literature,<sup>5</sup> in which a number between 10 and 15 participants per variable is suggested. Considering that the PSQI has seven variables, we initially calculated a minimum sample between 70 and 105 participants. However, as other authors<sup>2,3</sup> recommend that an exploratory analysis should have at least 100 participants, we decided to start the collection with 105 participants. We evaluated the KMO for these 105 participants. At that moment, we decided to double the sample and in the end, we managed to reach 209 participants. When analyzing the KMO value = 0.595, we proceeded with the analysis considering the literature recommendation<sup>3,5</sup> which determines that the minimum acceptable value is 0.5. In addition, it was important to review the KMO results. I realized that the value of 0.595 refers to the first two models that had seven variables. With the removal of a variable, obviously, the matrix changed and I only noticed it after the comment.

The KMO value for the model with six variables was 0.626, showing better adequacy of the sample in relation to the other models.

Regarding Bartlett's sphericity test, it is clear that it was an error when reporting the value of  $p=0.000$ . The correct value is  $p<0.001$ . It is important to highlight that the sphericity test and the KMO tend to be uniform and accepted the possibility of factoring the data matrix.<sup>6</sup> The results found are presented in the study tables, considering the analyses performed. The KMO and sphericity tests indicated that the models met the minimum criteria. The results are described and allow readers to make their judgment and draw their conclusions.

Our first objective was to assess the PSQI's reliability. This questionnaire, at first, was not prepared considering dimensions. Thus, we assessed the internal consistency and reproducibility of the general score. We considered that it would not be appropriate to assess the internal consistency among a few items, in each factor. Although we did not insert it in the text, we calculated the internal consistency for each factor, and as observed with the ICC values, the values were very close to the general value (0.70–0.75). When analyzing the reliability of the total PSQI score, we observed a moderate value (ICC 0.59–0.81). Most of the studies that assessed the PSQI's reliability were restricted to the analysis of ICC values, we decided that it was necessary to better explore these data. We compared the average values and detected significant differences. Therefore, we performed the Bland–Altman plot and managed to observe the presence of a systematic error. The positive bias with most participants above the zero line suggests that the score decreased in the second assessment. We consider it necessary to show readers that the measure has moderate reliability that can be influenced by random and systematic errors. We decided to calculate the Standard Error of the Measure and the Minimum Detectable Change to demonstrate the absolute error of the measure and what minimum change value should be considered. The results of the ICC indicate moderate levels of reliability. All measuring instruments are not error-free. Our results make it clear to the reader what the PSQI's level of uncertainty is regarding the reproducibility of the data.

Regarding confirmatory factor analysis, we followed the recommended procedures, based on theoretical concepts. This analysis aims to determine how a tested model can fit with the previously analyzed data,<sup>7</sup> and we specify in advance the structures that should be tested. The SRMR, RMSEA, and CFI values confirm that the models were valid. Among the tested models, we present the one that obtained the best performance in the aforementioned tests (Table 1) according to COSMIN guidelines (Consensus-based Standards for the selection of health Measurement Instruments).<sup>8</sup> The correlation observed between the items and showed in the figure can be explained from a theoretical point of view and has been observed in studies before and after ours.

Finally, I would like to highlight the importance of the observations made by the author, as many will contribute to future work. Our considerations about the internal consistency, reliability, and validity of the factor analyses are anchored in the adequate statistical tests and with interpretation consistent with the indexes established in the

**Table 1** Goodness-of-fit indices for the factor models of PSQI.

Model	SRMR	RMSEA	CFI	TLI	$\chi^2$ p-value
3-Factor	0.070	0.079	0.950	0.916	0.055
2-Factor	0.068	0.063	0.971	0.950	0.088
2-Factor (6 items)	0.060	0.060	0.975	0.973	0.246

PSQI, Pittsburgh Sleep Quality Index; CFI, comparative fit index; TLI, Tucker–Lewis index; RMSEA, root mean square error of approximation; SRMR, standardized root mean square residual.

literature. We present several analyses to allow readers to better understand and use the instrument. Regarding the factors, we seek to analyze three models that have been tested in different versions and it is indicated that the model with 2 factors and removing the “use of medicines” item presented the best fit. We emphasize that this model seemed to be more suitable among those tested. PSQI has proven to be a valid tool for different populations<sup>9</sup> and has been one of the main indirect measures of sleep quality. Regarding the highlighted phrase, it is important to note that we refer to the original version and that among the tested models of the Brazilian version of PSQI in our study, a specific model demonstrated a better fit. Our results are presented and interpreted in the light of the indices described in the literature, and do not even differ from the findings of two systematic reviews<sup>9,10</sup> on the subject. With a detailed account of the procedures and results, each reader can make their judgment and draw their own conclusions.

We never had the objective of torturing the data to obtain any kind of result, let alone because we have no conflict of interest. Our goal was to test the characteristics of an instrument and deliver its results to the academic and clinical community. If there was any intention of manipulating the result, it would not allow critical analysis and there would be no concern with a detailed inclusion of the results obtained. I conclude by thanking the opportunity for the debate. I believe it is clear, as well as in several areas of knowledge, that there are divergences in the literature regarding methods and their interpretations. Readers and researchers have access to the manuscript and will now have the position of the letter’s author and our response. It will be up to them to interpret in the way they consider most appropriate and to apply in future studies the path they deem most correct to seek the advancement of knowledge.

## Conflicts of interest

The author declares no conflicts of interest.

## References

1. Santos MS. 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. *J Pediatr (Rio J)*. 10.1016/j.jped.2020.11.003.
2. Tabachnick BG, Fidell LS. *Using Multivariate Statistics*. 7th ed. Boston: Pearson; 2018.
3. Hair JF, Black WC, Babin BJ. *Análise multivariada de dados*. 5th ed. Porto Alegre: Bookman; 2005, 593 p.
4. Fávero LPL, Belfiore PP, Silva FL, Chan BL. *Análise de dados: modelagem multivariada para tomada de decisões*. Elsevier; 2009.
5. Field A. *Descobrimos a Estatística Usando o SPSS*. 2nd ed. São Paulo: Artmed; 2009, 688 p.
6. Dziuban CD, Shirkey EC. When is a correlation matrix appropriate for factor analysis? Some decision rules. *Psychol Bull*. 1974;81:358–61.
7. Grimm LG, Yarnold PR. *Reading and understanding more multivariate statistics*. Washington, DC: APA; 2000, 437 p.
8. De Vet HCW, Terwee CB, Mokkink LB, Knol DL. *Measurement in medicine—a practical guide*. New York: U.S.A: Cambridge University Press; 2011.
9. Mollayeva T, Thurairajah P, Burton K, Mollayeva S, Shapiro CM, Colantonio A. The Pittsburgh sleep quality index as a screening tool for sleep dysfunction in clinical and non-clinical samples: a systematic review and meta-analysis. *Sleep Med Rev*. 2016;25:52–73.
10. Manzar MD, BaHammam AS, Hameed UA, Spence DW, Pandi-Perumal SR, Moscovitch A, et al. Dimensionality of the Pittsburgh Sleep Quality Index: a systematic review. *Health Qual Life Outcomes*. 2018;16:89.

Rodrigo Cappato de Araújo 

*Universidade de Pernambuco (UPE), Programa de Pós-Graduação em Reabilitação e Desempenho Funcional (PPGRDF), Petrolina, PE, Brazil*

*E-mail: rodrigo.cappato@upe.br*

<https://doi.org/10.1016/j.jped.2020.12.001>