

Peer Review Comments

Article: Condon, D M and Revelle, M 2015 Selected Personality Data from the SAPA-Project: On the Structure of Phrased Self-report Items. *Journal of Open Psychology Data* 3:e6, DOI: <http://dx.doi.org/10.5334/jopd.a1>

Article submitted: 21 April 2015

Article accepted: 09 July 2015

Article published: 17 August 2015

Editor: Jelte M. Wicherts

Responses for Version 1

Reviewer A: Name not given

Review Completed: 21 May 2015

This is an exciting dataset for anyone interested in personality or measurement. In addition to potentially interesting manuscripts focused on personality structure, as mentioned in the Reuse Potential, these data could be used to help someone new to the personality field gain a deeper understanding about different ways personality can be measured and how the measures related to each other.

The data could also be used to study methodological questions, such as item response patterns within and across measures. In addition, the data could be used for examples and homework assignments in statistical courses.

Thus, these data could potentially be used for publications, teaching, and to extend personal knowledge.

Despite having an R for Dummies book sitting on my Kindle, I have no experience with R; thus, I was unable to review the data documentation.

Reviewer B: Name not given

Review Completed: 11 May 2015

Review of JOPD manuscript, "Selected personality data from the SAPA-Project: On the structure of phrased self-report items"

This dataset from the Synthetic Aperture Personality Assessment (SAPA) project includes responses to self-report personality items by nearly 24,000 persons, each of whom provided responses to a subset of items (typically about 70) from a total pool of nearly 700 items.

In my opinion, this self-report dataset is potentially very useful for researchers who are interested in the relations among the many personality items, representing scales of several different instruments, that were included in this project. It will also be a useful dataset for researchers (and for course instructors) who would like a large dataset for application of various techniques in item analysis, in scale construction, and so on.

I am not at all an expert on the handling of missing data, particularly "massively missing" data of the kind reported, but it is reassuring that there were typically about 500 respondents who were administered both items of a given pair of items. This allows some confidence that the correlations among items will be reasonably close to the population values (at least for the population of persons who complete online personality assessments).

The authors have properly noted that their variable set would not necessarily constitute a representative sampling of the personality domain. As they stated, "...nor should it be expected that these items represent an unbiased or representative snapshot of human personality; the structure of these items will, to some extent, reflect the shared characteristics of the scales from which they were taken."

It could be useful if the data could also be provided in some additional format(s), such as Excel or SPSS, for the benefit of researchers and instructors who have not yet adopted R.