Reviewer’s report

Title: Reliability and Validity of the German Version of the Structured Interview of Personality Organization (STIPO)

Version: 2 Date: 19 July 2013

Reviewer: Kenneth Critchfield

Reviewer’s report:

The authors have done a fine job of responding to concerns I had on the previous review.

A few additional items remain, but are relatively minor and if addressed by the authors should strengthen the paper.

• Major Compulsory Revisions

Method Section:

(1) Re description of sample: Need to describe nature of the sampled settings in better detail to get a sense of the range of disorders present (e.g., how are patients typically referred, voluntary?, mandated?, some particular theme like suicidality?) and motivation for participating (e.g., it is not clear how participants were recruited for the study, was it part of clinic procedure for all who entered? Only those who chose? Were there incentives for participation?). Once concern, of course, is that you have sufficient expectation of heterogeneity on the underlying dimension of personality organization to answer your primary questions. Justification for this could simply read “a wide range of clinical problems were sampled from X to Y” (and then this is backed up by diagnostic descriptive statistics). A second issue is to allow future researchers to relate findings collected elsewhere under different conditions.

(2) Re: inclusion of ADP-IV. Need more clarity about rationale for inclusion of this instrument. Does it add something unique given that the SCID is also used? Perhaps this would be more compelling if the last sentence emphasized NOT that it is useful “for screening purposes” (it is a true statement, but in this context why “screen” when you have a full SCID assessment?). After mentioning the ADP-IV dimensional sensitivity, could you fairly argue that it is valuable to have the ADP-IV dimensions available for analyses rather than only the categorical distinctions provided by the SCID? Or perhaps emphasize that rater perspective (patient vs. clinician) might also be important? Use of a duplicate DSM-like method begs the question as to why. Please clarify in terms of how it will inform STIPO validity analyses beyond SCID.

(3) Regarding specific hypotheses: These hypotheses only identify areas where convergence is expected, but not where discriminant evidence is expected. Absent under Method (the authors do mention it later, in Discussion) is any clear statement that all these domains are expected (by theory) to inter-correlate to
some degree (i.e., that there is a unifying, underlying construct called “personality organization”).

I applaud inclusion of evidence that STIPO can discriminate among cases based on severity. This helps bolster the validity evidence. I also like newly included use of GAF to help make a clearer distinction between personality organization and general level of functioning.

• Minor Essential Revisions

(4) STIPO description, rationale, translation and scoring details: For the most part nicely handled. However, I request that you alert readers (in whatever way) that all of the STIPO domains are scored so that higher numbers indicate WORSE functioning in each area. The labels for these scales can be misleading to readers not already familiar with the instrument (e.g., “Identity” = “Identity diffusion/problems”, “Object Relations” = “problems with object relations”, “moral values” = “poor moral values,” “reality testing” = “poor reality testing”).

(5) Reliability and consistency information has been left out for SCID and OPD, please include

• Discretionary Revisions

(6) Results: I’m ok with the presentation of a raw correlation matrix, as they authors have done in the latest revision. However, it is not the strongest case that could be made to advance STIPO research. In order to find clearer evidence of specificity across instruments, the authors may wish to consider using regression methods to control for other STIPO scales when comparing across instruments. For example, partial correlation of STIPO Identity and BPI Identity Diffusion after controlling for the other STIPO scales may control the shared personality organization variance and allow the unique variance to shine through --- if successful, it’s more clear validity evidence; and if it doesn’t shine through, then you will be able to chart areas for further refinement.

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

Quality of written English: Acceptable

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

Declaration of competing interests:

I declare that I have no competing interests