Reviewer's report

Title: Psychometric evaluation of the community reintegration of injured service members computer adaptive test (CRIS-CAT)

Version: 2 Date: 17 August 2011

Reviewer: Morten Aa. Petersen

Reviewer's report:

The manuscript aims at assessing the reliability, validity and response burden of the so-called CRIS-CAT instrument. Such evaluations are essential when developing and testing new instruments. Therefore, the purpose of the manuscript is valid. The methods applied are well-known standard methods; no new, innovative methods are applied. Hence, the manuscript seems relevant for the potential users of the CRIS-CAT, but likely less relevant for readers more generally interested in developments in psychometric methods and instrument validation.

Generally the manuscript is clear and well-written and the methods applied are appropriate and well described (however, see comments below).

Major Compulsory Revisions

1. Methods, Data Collection, Study Measures: “At Visit 1 we collected data on other measures which we believed would be correlated with CRIS-CAT scores if the instruments were measuring community integration...” Given that almost anything correlates and that the set of measures used for these correlations are not measures of community integration, this seems as a rather vague hypothesis and as rather vague evidence that the CRIS-CAT measures community integration. For example, PCS and MCS may be significantly correlated in many samples, but this seems as poor evidence that PCS measures physical performance and MCS mental. Ideally, one would use some kind of “gold standard” for such correlations, however, if such is not available (which is rarely the case), then I think the authors at least should set up some more specific hypothesis of what would be expected of a measure of community integration. For example, how do they expect the CRIS-CAT measures to correlate with each of the other measures (high, moderate, or low correlation), do they expect the three CRIS-CAT scales to correlate similarly with the other measures? Otherwise, I think the conclusions regarding (concurrent) validity should be toned down.

Minor Essential Revisions

1. Methods, 3rd paragraph, line 4: “In previous work we developed the CRIS CAT by...” No references are given to this previous work. If this work is published please provide references.

2. Results, Cohort Study Sample & Table 2: There were several findings of
significant differences between the participants in the cohort study and those lost to follow-up. Hence, the participants seem as a selected, non-representative sample of OEF/OIF veterans. How does this affect the generalisability of the findings?

3. Figure 4: What is meant by RRR on the x-axis?

4. Discussion, 5th paragraph, line 2: “...can be administered with minimal respondent burden.” Given the quite limited evaluations of response burden (only the number of items required for one CAT setting is reported, no response time, patient feedback etc. are reported), this seems as quite a strong statement. This needs further justification or should be toned down.

Discretionary Revisions

1. Methods, Statistical Analyses, 4th paragraph, line 6: “We ran 3 separate multinomial regression models to predict change in...” Why three separate regression and not one multiple regression including all three CRIS-CAT scales? This would elucidate which CRIS-CAT scale might be the strongest predictor.

2. Table 2: It would be more informative if the actual p-values, and not just *’s were shown (this is done in table 3), possibly highlighting significant findings, e.g. p-values in bold.

3. Table 5: Assuming that the SF12 v1 measure used is PCS v1 when predicting PCS v2 and MCS v1 for MCS v2, it is surprising to me that in all but one case, the CRIS-CAT scales at visit 1 seem as better predictors of PCS/MCS v2 than PCS/MCS v1. Was this really the case? If so, please try to explain/comment on this.

Minor issues not for publication

1. Methods, line 2: “of OEF/OIF veterans” is written twice

2. Methods, line 7: “examination the CRIS-CAT’s”, missing “of” after examination

3. Statistical Analyses, 2nd paragraph, last line: “1/(1+(standard error)2 )[37]” seems to refer to a reference 37, but there are only 22 references in the list of references.

4. Discussion, 6th paragraph, line 6: “overestimate these correlations. [30]” seems to refer to a reference 30, but there are only 22 references in the list of references.

5. Figure 2a. “Reliability of CAT estimates: Extent Scale” is labeled 1a.

6. Figure 2a-c: Cannot read the titles for the y-axes.

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 declare that I have no competing interests