Reviewer's report

Title: Exploratory factor analysis of self-reported symptoms in a large, population-based military cohort

Version: 1 Date: 6 August 2010

Reviewer: William D Killgore

Reviewer's report:

Minor Essential Revisions

1) The authors mention that the demographics for this sample are reported elsewhere. My own preference would be to actually have at least some rudimentary description of the sample available when I am reading a paper, rather than have to track down a different paper in a different journal just to see basic demographics. I will defer to the editor on this, but perhaps even a line or two just to describe the basics of the sample. Readers could still be referred elsewhere for more detailed information.

2) On page 7, the authors describe that they factor analyzed data from a variety of smaller scales imbedded in the survey. For some scales, all items were analyzed, while for others, only some items (e.g., “most items” line 7) were analyzed. How was the inclusion or exclusion of items determined?

3) The authors point out that “an important finding of this study was that the majority of the factors appeared to load strongly based on how symptoms were grouped according to location within the survey.” However, this conclusion cannot be demonstrated conclusively. First, location is confounded with content. It may be that subjects responded to the items similarly because they were in the same location of the survey, as suggested by the authors, but it may just be that it was because all of those items queried similar content. It is impossible to tease that apart. Secondly, item content and location are also confounded with response format. The various items for a particular scale (e.g., anxiety, PTSD, somatic symptoms) had a common format (e.g., dichotomous, 3-item, 4-item, 6-item responses). Thus, questions that asked about depression often had 4-item response formats, while questions that asked about disordered eating had a dichotomous response format, and items that asked about vitality had a 6-item format. Thus, the factor structure may have been partly reflecting response format as well as content or location. This limitation needs to be addressed in the discussion section.

4) The study seems to be entirely descriptive. There is no theoretical rationale or models put forth and it is not entirely clear what the ultimate utility of such an exploratory factor analysis would be. At this point, it shows that the current survey includes items measuring a variety of constructs and they appear to hold together relatively well. Is this the ultimate goal, or were there any theoretical
bases for this analysis? Will the results of this analysis be used to modify the survey (i.e., dropping redundant items)?

**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