Reviewer's report

Title: Validation of the Finnish Version of the SCOFF Questionnaire among Young Adults Aged 20 to 35 Years.

Version: 1 Date: 24 July 2008

Reviewer: Jonathan Mond

Reviewer's report:

The authors examined the validity of a 5-item self-report measure of eating disorder symptoms, the SCOFF, in screening for cases of eating disorders in a general population sample of women and men aged 20-29 years (n=1316). Validity coefficients were calculated in a subgroup of participants (n=541) who completed a diagnostic interview (in addition to the SCOFF), of whom n=9 were identified as eating disorder cases. At the optimal cut-off of 2 or more positive responses, sensitivity (0.78) and specificity (0.88) were high, whereas positive predictive value (PPV) (0.10) was low. The authors conclude that, in general population samples, the SCOFF may be useful in ruling out cases of eating disorders, but that PPV is too low for the measure to be used to identify probable cases.

This is a useful contribution to the small literature concerning the utility of various self-report measures in screening for eating disorders in unselected samples. Attractive features of this manuscript are the use of a general population sample, comparatively large sample size, and identification of “true cases” on the basis of an established diagnostic interview. There are, however, some problematic aspects of the manuscript in its current form. In particular, important aspects of the study methods need to be clarified, the Introduction is unduly brief, and the Discussion is overly long (also page numbers would have been helpful). Specific points that need to be addressed, by section, are as follows:

Abstract
- It would be better to note the number of participants who completed questionnaires rather than (or in addition to) the number of individuals approached.
- It would also be helpful to note the number of cases on which the ROC analysis was based and the diagnoses.
- The figure of 8.3% for PPV does not accord with the figures given in the Results section.

Introduction
- A recent study by Mond and colleagues (Behaviour Research and Therapy 2008; 46: 612ff), in which the utility of the SCOFF in screening for eating disorders in primary care was compared with that of the EDE-Q, should be incorporated in the Introduction and/or Discussion. Many of the issues pertaining
to the use of such measures in unselected samples are discussed in detail in this manuscript. In particular, there is nothing new about the problem of low PPV when screening for eating (and other psychiatric) disorders in unselected samples.

- It seems odd to leave all mention of what previous studies have found concerning the validity of the SCOFF to the Discussion. By giving some account of the existing literature in the Introduction, the significance of the study would be clearer. As it is, the Introduction doesn’t give the reader much of a sense of why this research is useful.

Method and Results
- The sentence “Persons reporting … participate in the mental health interview” is repeated in the second and third paragraphs of the first page.
- At the end of the first section, where response at the first stage is detailed, it would be helpful to clarify how many of the 982 persons approached for interview were screen+ vs screen-.
- If there was no information about the characteristics of participants (n=547) who did not return completed questionnaires this should be stated and the implications of this considered in the Discussion. If there is information about age and gender of respondents vs non-respondents then this at least should be presented. The authors do consider potential bias resulting from non-participation at the second phase (interview) phase of the study, but that is a different issue.
- In the section on “mental health assessment”, it would be important to clarify the operational definition of EDNOS employed, given that there is no clear consensus on this, given that decisions made in this regard influence the nature and number of the cases identified, and given that the majority of cases of EDs in both community and clinical samples do not meet formal criteria for AN, BN or BED.
- I agree that current (rather than lifetime) cases are the appropriate group for the analysis, but the reason given for this (that cases were relatively few in number) doesn’t make sense (low numbers would be a reason to consider lifetime cases). The reason that current cases are the appropriate group is that it is of limited interest to know that someone may have been a case in the past.
- With respect to the items of the SCOFF, would the authors expect any problems of translation? For example, there is scope for ambiguity concerning the item “do you make yourself sick …”, which could be interpreted to mean feeling sick on account of having eating too much (rather than using self-induced vomiting as a means of weight control) if care was not taken.
- It would make more sense to cover all material relating to the recruitment of participants (i.e. including the first paragraph of the current Results section and details of cases and non-cases included in the ROC analysis) in the Method section, leaving just details of validity coefficients/ROC analysis for the Results.
- The authors note, in the first paragraph of the Results, that mean scores on the SCOFF did not differ significantly between participants and non-participants at the interview phase, but it appears that there was a trend towards scores being
higher among non-participants (p=0.11) and mean scores can conceal things. Were there differences between groups at the item level?

- Wherever details of cases are given, it would be helpful to know the diagnoses. Also, the prevalence of any ED diagnosis in women (9/900 = 1.0%) seems low, raising questions as to whether there might have been a response bias (at either phase) and/or overly restrictive operational criteria. I wonder whether it would be possible to reanalyze the data using a less stringent operational definition of EDNOS. A comparison of this kind would be of interest and might also permit analysis of validity coefficients by age group (cf. Mond et al, 2008).

- Table 2 could be deleted, as the results are essentially the same and are covered in the text.

- Table 3 also could be deleted, but if it is to be included then it should be included in the Introduction rather than the Discussion.

Discussion

- This section could be shortened considerably by moving some material to the Introduction and eliminating repetition and material that will be self-evident to most readers. For example, the final paragraph of the second to last page could be deleted without any great loss of substance and the final two paragraphs of the paper could be replaced with the one brief conclusion. I suggest the authors review this section in order to (a) decide what material would better be included in the Introduction and (b) eliminate repetition and other redundancy.

- The authors draw attention to the high PPV observed in the study of Parker et al and note that this likely reflects higher prevalence in this and other studies. Differences in prevalence are certainly one factor, but differences in study methods and operational criteria for diagnosis also need to be considered. From memory, in the Parker et al study, eating disorder “cases” were identified on the basis of a second self-report measure (the EDE-Q?). For this reason alone the findings need to be interpreted with caution. In practice, it is likely that the validity of the SCOFF in identifying primarily sub-threshold cases is what was tested in the Parker et al study, which incidentally is consistent with the strategy of a staged approach to screening subsequently mentioned by the authors.

- The authors suggest that the SCID is the “golden (gold?) standard” for identification of eating disorders, but I think most in the eating disorders field, if forced to nominate such a thing, would give the title to the Eating Disorder Examination rather than the SCID.

**Level of interest:** An article of importance in its field

**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.