Reviewer's report

Title: Measuring bias in self-reported survey research: validating the Crowne-Marlowe Social Desirability Scale in Ethiopia, Kenya, Mozambique, and Uganda

Version: 1 Date: 21 May 2011

Reviewer: Andrea Petróczi

Reviewer's report:

The aim of the paper is to provide a social desirability scale to complement HIV risk behaviour research relying on self-reports among African populations. Whilst this aim is well justified, the paper – as it stands – falls short of this promise. The comments below intend to help the authors to enhance their work to deliver on this worthy aim and produce a paper that warrants publication. On a positive note, the paper is clearly written and well structured. The translation process appears to be adequate and the sampling is sufficiently rigorous.

Major compulsory revisions:

1. Marlowe-Crowne SD scale is misinterpreted in the text. SD is a tendency of individuals to reply in a way that will be, presumably, viewed favorably by others (emphasis on the ‘individuals’). In other word, Marlowe-Crowne SD taps into a personality trait coined ‘approval seeking’ (Crowne & Marlowe 1964) or later, more appropriately ‘avoidance of disapproval’ (Crowne 1979). As such, SD effect individuals’ responses on other self-reports to a varying degree. Consequently, SD effect should be discussed in relation to a potential bias it may cause in other psychological/behavioural measures. Marlowe-Crowne SD scale measures individuals’ propensity for impression management. I strongly suggest reading McCrea & Costa 1983 (McCrea, RR & Costa PT: Social desirability scales: More substance than style. J Consulting Clin Psychol 1983, 51:882-888). Petroczi & Nepusz 2011 (Petroczi A & Nepusz T: Methodological considerations regarding response bias effect in substance use research: is correlation between the measured variables sufficient?, Subst Abuse Treatment Prev Policy 2011, 6:1) might be useful for an overview on the use of SD scales, along with Delroy Paulhus' work on SD in general.

2. Evidence for reliability (internal consistency) is provided as Cronbach alpha coefficients. The authors should justify this choice (Cronbach alpha is an internal consistency estimate for Likert (or Likert-type) scales and how it might affect the reliability estimates. Kuder-Richardson formulae (KR-20 or KR-21) are better reliability estimator for scales with dichotomous response options (e.g. true/false).

3. Further to evidence for reliability, I would have liked to see some attempt to explain the less than desirable Cronbach alpha values. Have the authors conducted any analysis (e.g. a simple item to total correlation or confirmatory
factor analysis) that might shed the light on this? If so, please include. If not, please do and include.

4. Despite the title and specific aim, the paper does not provide any evidence for validity. I suggest that the authors consult the APA Standards for Psychological Testing before addressing this point.

5. Discussion and conclusion must be re-written once the above points are addressed.

Minor essential revisions:

6. Justification should be provided for selecting the Marlowe-Crowne SD scale over other SD scales.

7. The short forms of the Marlowe-Crowne SD scale are notoriously unreliable. Why did the authors decide to translate and test them?

8. I would have liked to see a slightly more substantial discussion on social desirability.

9. Abstract: first line of results are not results. These are sample characteristics (thus belong to the method).

10. Methods: how were the data collected? Self-administered questionnaire or interview? A discussion on how the method might affect the responses should be included.

11. Excluded items: why was No 29 excluded?

Discretionary revisions:

12. Given the overarching aim of this study (which I agree with), I would have expected to see the scales included as appendix or at least to see some information on how the copy can be obtained.

**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.