Reviewer's report

Title: Scale development on consumer behavior toward counterfeit drugs in a developing country: A quantitative study exploiting the tools of an evolving paradigm

Version: 3 Date: 20 February 2013

Reviewer: Richard Fielding

Reviewer's report:

BMC Public Health: Scale development on consumer behavior toward counterfeit drugs in a developing country: A quantitative study exploiting the tools of an evolving paradigm

Counterfeit drug use is, as the Introduction of the ms states, an very serious international problem. Hence this ms addresses and important public health issue. It also has another strength in that it reports on a Sudanese Arabic-speaking population for whom available data is sparse relative to English-speaking populations. Hence the ms is potentially a useful addition to the international literature. However, I have a number of concerns about the methods used.

Title: (or key-words) should mention Arabic-speaking Sudanese population.

Introduction. Ok

Methods:

Sampling –

1. The sample was an “availability” sample. Explain what this means. How precisely were subjects identified and invited to participate? P8 para 2 states that people were sampled according to available demographic information. Again, are these the people most likely to be using counterfeit drugs? I would think not.

2. Given that the sample appears to be disproportionately highly educated, this would suggest that a non-representative group were captured. These people would be those most likely to have higher incomes and hence be better able to afford authentic drugs. I would therefore expect that their attitudes towards counterfeit drugs would differ from those who had little alternatives, but who constituted only 5% of the sample. How do the authors respond to this potentially very serious bias?

3. Generally, sampling details lack clarity and information needed to make informed decisions about the quality of the sample(s) obtained.

4. Were both samples recruited in exactly the same way? Where they comparable? Was sample 1 a subset of the overall sample? Please explain.

5. Sample size appears to have been adequate for the number of scale items examined.
6. Explain the factor selection criteria used in the EFA. Which rotation method was used and why? The results

Results
6. As a rule sample characteristics should be presented in Table 1.
7. The authors found 11 factors which they claim supports the “multidimensional nature of the scale”, but as far as I can see, the authors tried to measure just four dimensions “namely, attitude, subjective norms, motivation and behavioral intention.” (para 1 p12). Moreover, 11 dimensions seems a large number from 41 items. This suggests that the extraction criteria used were too loose, or the item elimination was insufficient. As some factors had nine items, that means the majority of factors have at most two two.

8. p11. “Reliability assessment” “Reliability assessment of the 10 constructs…” But earlier you say there were 44 items and 11 constructs. Now there seem to be 37 items and 10 constructs. What happened?
9. Query 7. above raises a second concern about the Cronbach alphas; alpha is an indicator of scalability. The internal consistency is a function of the diversity of weights for each item and the number of items. Alphas improve as the number of items in a scale decrease. With only two items, diversity is low, and unless item weights differ dramatically, alphas will be acceptable. This however, doesn’t tell us much as most factors have only two items.
10. The first line mentions results derived from CFA confirming convergent/divergent validity, but the previous section reported only EFA was used. Please explain.
11. Convergent and divergent validity referred to in this paper are for the factors in the instrument, not for the instrument as a whole, for which these measures have not been ascertained. Please make this explicit.
12. What is the total variance accounted for? How is that variance apportioned through the 10 sub-scales?

Discussion
13. There is no attempt to relate the 10 or 11 subscales to the TPB which is used to justify the instrument development. There is no discussion of what those subscales reflect and how these relate to the TPB, or consumer behaviour. For example “awareness of social consequences” relates to the TPB how, exactly?

Minor points:
P8 para 2: “the authors nonetheless attempted to be representative”. This sentence needs rephrasing. It is not the authors who need to be representative, but they do need to ensure that the sample is representative.
P10 “privately employed”. Do the authors mean “self-employed”?
P10. Which software and programme was used to perform the FA and EF? Which type of factor analysis was used? I notice at the foot of Table 2 that Varimax rotation was used, rather than the recommended Oblim for scale development. Please justify.
Table 2: no probability is zero. Report p values as “p<0.001” rather than “0.000”.

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

**Quality of written English:** Not suitable for publication unless extensively edited

**Statistical review:** Yes, and I have assessed the statistics in my report.

**Declaration of competing interests:**

I declare I have no competing interests